



# Biotechnology software in the digital age: are you winning?

Cornelia Johanna Franziska Scheitz<sup>1</sup> · Lawrence J. Peck<sup>1</sup> · Eli S. Groban<sup>1</sup>

Received: 7 November 2017 / Accepted: 9 January 2018 / Published online: 16 January 2018  
© The Author(s) 2018. This article is an open access publication

## Abstract

There is a digital revolution taking place and biotechnology companies are slow to adapt. Many pharmaceutical, biotechnology, and industrial bio-production companies believe that software must be developed and maintained in-house and that data are more secure on internal servers than on the cloud. In fact, most companies in this space continue to employ large IT and software teams and acquire computational infrastructure in the form of in-house servers. This is due to a fear of the cloud not sufficiently protecting in-house resources and the belief that their software is valuable IP. Over the next decade, the ability to quickly adapt to changing market conditions, with agile software teams, will quickly become a compelling competitive advantage. Biotechnology companies that do not adopt the new regime may lose on key business metrics such as return on invested capital, revenue, profitability, and eventually market share.

**Keywords** Biological design · Cloud compute · Cloud security · Computer-aided design · Digital

## Abbreviations

CAM	Computer-aided manufacturing
CAD	Computer-aided design
ROIC	Return on invested capital
IT	Information technology
IP	Intellectual property
IPO	Initial public offering
API	Application programming interface

## Introduction

Biotechnology is an industry prime for disruption for the following reasons: (i) it is rapidly growing [8], there is frequent IPO activity [18], and takeovers are plentiful [11], making market entrance for small start-ups with venture capital backing a high probability; (ii) the tools and techniques for biological manufacturing are advancing quickly, with the price of reading and writing DNA falling faster than Moore's law predicts, leading to the democratization of information and materials [4]; (iii) rising levels of investment in biology bring the size of Series B rounds to over \$50 M [10] and allow smaller companies to enter the space with fully automated laboratories and production capabilities, making

automation a commodity rather than a competitive advantage; and (iv) increasing availability of affordable infinite computing capacity removes the requirement for expensive, on-site infrastructure [19].

Software in Research and Development groups is lagging behind the current technology landscape. As an industry, we employ amazingly smart people, pay them a lot of money, and then provide poor tools for them to accomplish their design work [5]. Many times, the applications a scientist works with during the workday are far less advanced, far less secure, and deliver an inferior user experience compared to the applications this same scientist will use on a smartphone after work. Why do we pay so little attention to ROIC or business efficiency when judging the software that these scientists use to do their jobs? There are a number of reasons for this. First, design is still believed to be something that only a scientist can do. Therefore, there is little incentive to provide software that thinks with the scientist, instead of the one-way Human–Computer Interaction we have today. Second, there is entrenched thought that biotech, genetic engineering, synthetic biology, whatever we want to call it, requires unique design tools and software capabilities compared to other markets. As such, in-house software is seen as a competitive advantage. Third, cloud services are still not fully adopted, and while certainly more secure than any local storage, many still believe that local storage of data is more secure than the cloud. Together, these factors compel biotechnology companies to maintain large, internal IT teams

✉ Eli S. Groban  
egroban@gmail.com

<sup>1</sup> Autodesk Life Sciences, San Francisco, CA 94111, USA

and spend both capital and operation expense on in-house servers and in-house software development, decreasing the company's ROIC [3]. Moreover, when evaluating new software tools, many of which seem expensive or less secure, we may not rationally consider the fully loaded costs of our in-house software teams or hardware systems, thus ignoring business efficiency.

Operations in most large and many small Biotechnology companies are a fully automated pipeline, and computer-aided manufacturing (CAM) for biology is mostly solved. The level of automation in the typical pharma laboratory parallels advanced manufacturing facilities such as those in the automotive world, and they benefit from software advances in the CAM world over the past 40 years. Instead of having to develop software in-house, companies use off-the-shelf design and infrastructure management software for their automated laboratory designs as well as their factories and production centers. Since early stage biology discovery companies are securing Series B rounds over \$50 M, they can also enter the space with fully automated laboratories and production capabilities [10]. Over the past few decades, automation in life sciences has transitioned from a competitive advantage to an affordable standard. Design, however, is not as advanced. It is still seen as a competitive advantage, one that relies on in-house software and algorithms. Moreover, software for design appears to require specialized tools and algorithms for biologists. As biologists, we think that our design process is unique.

### **Biology is not the first industry to cross the digitization chasm**

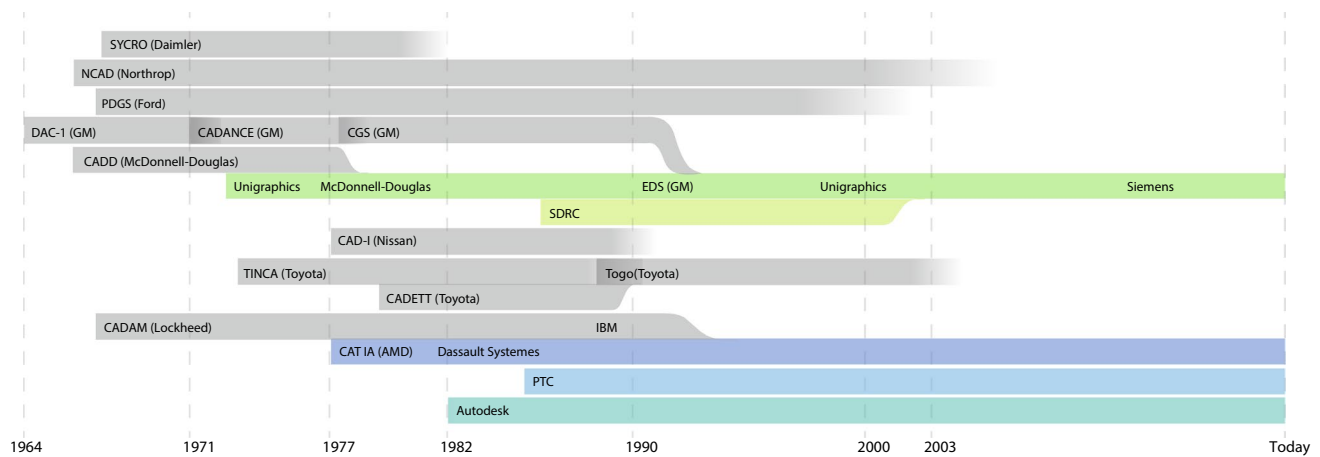
While the industry is happy to purchase software to design and operate production centers, this is far from the truth for designing the biological product itself, be it antibodies or strains of yeast. Here, the state of the art is to rely heavily on the scientists with a light assist from software to track data. In cases where efforts have been made to automate progress, such as at Amyris, most solutions are developed in-house. In fact, the current environment of biological computer-aided design (CAD) software bears a striking similarity to the physical CAD software landscape in the 1970s [2]. As early as the mid 1960s and all through the 1970s and 1980s, large manufacturing firms were investing in software development groups to build their own proprietary design software. Among the automotive companies, Ford was investing in PDGS, General Motors built CADANCE and CGS, Mercedes-Benz had SYRICO, Nissan wrote CAD-I, and Toyota made both TINCA and CADETT. The aerospace manufacturers also threw resources into design tools with Lockheed producing CADAM, McDonnell-Douglas making CADD, and Northrop developing NCAD. Auto and airplane

design was challenging, so this proprietary software ran on dedicated high-powered computational resources housed at each specific company. Manufacturers saw their homegrown tools as key intellectual property, heavily investing in unique algorithms and mathematic solutions which they perceived as their edge over the competition. Purpose built curve and surface types such as Cinci parabolas, Bézier curves, Coons patches, and Gordon surfaces have all given way to commercially available software solutions produced by the four main providers of manufacturing CAD/CAM software: Siemens NX, PTC, Autodesk, and Dassault Systemes (Fig. 1). It is of note that most in-house developed packages were never integrated into a commercial suite, but in fact ultimately required expensive and time-consuming replacement and data migration initiatives. Automotive and aerospace manufacturers recognized that they were better served by focusing their resources on their core competency, autos, and airplanes while relying on professional software firms to deliver world class software tools.

### **Technology advances are driving changes in design paradigms**

The way that we design in the physical world is changing. 50 years ago, when design software first hit the market, the interaction model between the designer and the computer was unidirectional. Designs appeared intuitive, defined by human comprehension and design ability. What appears to be strength on first sight was also a weakness though as designs were limited. However, at that time, the computational power was also quite limited, unable to design without directive input from the human. Software in the design space served as a digital replacement within an analog process. For example, using an excel sheet to compute rather than a pen and paper, or a calculator. It serves as a time saver, but not a game changer. The availability of cheap and infinite compute in the 21st century allows for the introduction of technologies powered by machine learning. These new technologies change the human computer interaction model, moving from a one-sided approach towards two-way communication. A human makes suggestions, and the computer can process and return these, while at the same time making suggestions as well. This requires leveraging the power of infinite compute and is a step change that redefines the way in which we approach design and Human–Computer Interaction.

The way that we design in the Life Sciences is changing in the same ways as in the physical world. Life Sciences companies are fully automated machines, and as described above, CAM for biology is a solved problem. CAD, however, is believed to be something that only a scientist can do. This is rapidly changing. There are two examples over the past few years, where computers played a significant role in the



**Fig. 1** CAD/CAM packages developed by companies in the automotive and aerospace industry (grey bars) and software companies (color) over time. The individual bars are labeled with the tool name and the company name in parentheses for all non-commercial tools. Otherwise, the bar is labeled with the company name. Gradients indicate when tools were combined into new packages or when they were deprecated. Multiple company names on a bar indicate changes in ownership. While over 14 independent packages were developed,

with tremendous amounts of man hours dedicated to each one, today, only four packages still exist. Two of those have been developed independently from the automotive and aerospace industry. Six of the nine in-house developed packages (grey) were never integrated into a commercial suite, but in fact ultimately required expensive and time-consuming replacement and data migration initiatives. Today, all automotive and aerospace companies use a combination of the four available commercial tools with an additional layer of internal IP

design process for two separate drugs. The first example is a drug for liver disease from a company called Nimbus Therapeutics in Cambridge, MA. They developed a drug molecule leveraging computation and sold this molecule to Gilead for \$400 M, with further royalties up to \$800 M [6]. The second example is one of a professor at the Vancouver Prostate Cancer Center, an affiliate of the University of British Columbia [16]. Using computation, this professor developed a drug for advanced refractory prostate cancer that the University sold to Roche for \$140 M thus far. Let us keep in mind that while these advances were made possible by computation, they did not leverage advanced tools such as generative design or AI that are becoming commonplace in the design of the physical world. Therefore, the future seems bright, but it begs the question: How can we leverage these new tools and technologies in biotech? We would like to argue that the prerequisite is for companies to become completely digital—or in other words data driven.

## Towards data-driven biotechnology

Digital solutions are developed specifically to save time, reduce costs, streamline complex processes, reduce error, and create agility in resources [3]. There are few life science companies that are true digital companies, having fully digital, data-driven processes. Just collecting copious amounts of data, to supply offline to subjective scientists does not equate to data-driven process. Data driven means automatically collecting, mining, and leveraging data to

guide future decisions. At the same time, there are many life science companies that believe that they are digital and are fully automated. These companies usually have some digital solutions patching analog processes. For example, downloading a CSV file to control a liquid handler is not digital. Even though the files themselves are generated by a computer and used by a robot, creating and downloading files brings certain sources of error such as versioning and data loss. Communicating with a liquid handling robot via an API connection is a digital process, not having to create files and move them among different places. All companies should strive to be digital. Most life science companies are having active discussions around digitization, cloud computing, and collaboration. The differences lie in how they are addressing these concerns. Their realities fall into three different categories ranging from full adoption to professed adoption to scared to adopt. In the life sciences, less than 10% of companies fall into the first category, and most are split between categories 2 and 3 [3]:

## Full adoption

These companies truly understand how to drive value with digital transformation [3]. They have a transformative vision on how digital can be used to drive innovation. They are coordinating digital directives across the company and develop a digital culture that can quickly react to changing market conditions or trends. They embrace cloud services provided by companies such as Amazon, Google, Microsoft, or Oracle, realizing that these solutions are more secure than

their internal networks [14]. Very few companies fall under this category. This usually falls under the realm of small start-ups. Start-ups in this field tend to focus on one area of the design process as the entire process is too large for one small company to address.

### Professed adoption

Information technology teams at various companies hear from internal customers terms like “cloud” and “digital design”. They implement home grown solutions that address each of these concerns, but these solutions are neither state of the art nor user friendly [3]. Usually, both customer groups (management and scientific users) believe that their internal IT solutions are state of the art, or are at least as close as they can be with current fears around security. In reality, these solutions are often clunky, time consuming to use, and are not as secure as an external software vendor would be. There are several prominent examples, where large pharma or biotechnology companies have been hacked. The most recent and prominent was the 2-day shut-down of Merck US due to ransomware that prevented any of the employees at the company from accessing their computers [15]. Other prominent examples include hackers accessing Bayer [7], Boston Scientific, Medtronic, St. Jude [9], and three unnamed major European pharmaceutical companies [17]. This is a small list compared to the larger list of companies hacked since the AOL data breach of 2004 [12].

### Scared to adopt

Many biotechnology and pharma companies will not even entertain the concept of the cloud, although this is changing. They believe that there may be benefits to cloud software and cloud technologies, but they are not convinced that the benefits outweigh the perceived risks. On premise, storage is perceived to be the safer option, regardless of the reality that on premise carries a higher risk than data storage with certified cloud providers. Publicized break-ins and data security breaches drive internal fear, rather than driving change. The internal teams are also unable to add security without sacrificing features, meaning that the internal systems become slow, hard to use, and do not scale, resulting in productivity losses.

It is a given that digital transformation is going to happen across all industries, including biotech. Ultimately, the question comes down to: What does this mean for my business? Other markets such as banking, insurance, travel and hospitality, and, of course, high-tech, have already faced this challenge and we can extrapolate some information. Across all these industries, research shows companies that strategically digitized show increased revenue, profit, and market capitalization [3]. However,

furthermore, companies with professed adoption, while having increased revenue, show decreased profit and market capitalization and are overall on par with companies that have not adopted digital at all. Finally, digitization provides one other major advantage—the agility to react quickly to market and business changes. There are two takeaways for companies adopting fully data-driven processes: (i) a company needs to be all-in to get it right and (ii) companies that get it right it will reap exponential improvements across the board [3].

### What do the winners look like?

Determining how your company is responding to the current environment, and whether they are putting together a winning strategy, is not easy. While initiatives need to be chosen strategically to fit the company and the business problem at hand, most processes within a company can be improved by digitization and should be evaluated. This spans from early R&D to pipeline and supply management. Some of the behaviors are compiled in Table 1. This should help to determine whether your company is a leader in the global race towards digitization.

Many of the examples of complete digitization lie outside of life sciences as we have only begun this journey in the field of biology and biotechnology. A great example of a complete digital transformation is Nike [13]. Not only have they optimized their sales in existing markets, they have adjusted their supply chain, revolutionized their internal design pipeline, and moved closer to the customer. For Nike, the sale used to be the end of their relationship with the customer. By moving to digital, the sale was only the beginning of this relationship. Innovative digital products such as a sports watch allowed Nike to deliver content to customers rather than advertisements. Location-specific grids and virtual reality city tours enhanced the running experience and placed the Nike brand at the center of the customer running experience. Customers now interact with Nike long past the initial sale. Within life sciences, AstraZeneca is a great example of a company that recently laid the foundation for full digital adoption [1]. Over the past few years, this company moved from legacy tools to up-to-date cloud-based software products. In addition, they now leverage not only software as a service, but hardware as a service as well, allowing their scientists access to compute via Amazon Web Services. At AstraZeneca, not only has the digital transformation already saved them hundreds of millions of dollars, it makes their employees more productive, consolidates and secures their key intellectual property, and enables them to re-center their business strategy around agile, rapid innovation.

**Table 1** Example behaviors, broken out into function, for those companies embracing digital transformation and those companies content with the status quo

Specific Function	Transformative	Sedentary
Infrastructure	Divest from expensive and hard to maintain internal infrastructure. Leverage increased computational power	Continually invest in internal servers as computational power increases
IT team	Repurpose IT teams from supporting legacy hardware to enabling cloud services	Continue to employ internal resources to maintain legacy infrastructure
Security	Prevent data loss by leveraging secure platforms such as Amazon Web Services, Microsoft Azure, or Google Cloud. Understanding that encryption and obfuscation across servers adds to already robust security	Remain on internal servers, exposing company data to hacks and theft. Continued belief that data behind a corporate firewall is safer than on heavily protected cloud platforms
IT reporting	Bill IT resources directly back to R&D program budgets, allow R&D teams to evaluate IT performance, enable transparency in IT spend	IT continues to operate independent from business units, not a service business, no transparency in cost/benefit
Data/learning	Centralize and secure all company data to allow progress via machine learning or artificial intelligence	Allow scientists to work in off line legacy systems that have an inability to aggregate data for learning
Cost	Conduct objective cost/benefit and ROIC analysis on employing internal resources to write software vs. investing in external software	Shy away from seemingly high cost, advanced cloud-based software platforms by focusing on the short term, instead of analyzing long term savings employed by resource optimization
Collaboration	Switch to software platforms built for cloud-based technologies that enable seamless communication and collaboration among all scientists	Continue to encourage desktop software platforms, unable to leverage cloud compute or enable collaboration

## Conclusion

The main disruptive force in biology, just as in all industries, over the next decade will be software. Biology companies, however, are ill suited to respond to this change. Their success comes from a slow and steady, conservative approach. In this new age of digitalization, where one must adapt quickly, with nimble and fast operational software teams, this strength becomes a huge weakness. Looking across industries, companies that embrace digital and cloud technologies outperform their conservative competitors on various metrics such as ROIC, revenue, and profit. This trend will certainly not bypass the life sciences. As an industry, we need to take an objective view of the cloud, before our data are stolen, and realize that we are not software companies. Just like we would not look to Toyota to supply our physical CAD software, we probably should not look to Pfizer to supply our biological CAD software. While internal software and algorithms may work for now, we need to objectively look at the opportunity cost of these efforts and the subsequent loss in ROIC. Companies that can do this objectively will end up at the top, while those that cannot may end up with data leaks, massive, and costly software teams, frustrated scientists, and an upper management team wondering how they lost.

**Acknowledgements** We thank Roger Mollon for fact checking our distant recollections of the history of CAD.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Bloomberg J (2016) Cloud-Centered IT transformation at AstraZeneca. In: Forbes. <https://www.forbes.com/sites/jasonbloomberg/2016/10/20/cloud-centered-it-transformation-at-astrazeneca/#6beaeabc4cd7>
2. CAD software— history of CAD CAM. <http://www.cadazz.com/index.htm>
3. Capgemini Consulting (2012) The digital advantage: How digital leaders outperform their peers in every industry. MIT Sloan Manag Rev 1–24. [https://www.capgemini.com/wp-content/uploads/2017/07/The\\_Digital\\_Advantage\\_How\\_Digital\\_Leaders\\_Outperform\\_their\\_Peers\\_in\\_Every\\_Industry.pdf](https://www.capgemini.com/wp-content/uploads/2017/07/The_Digital_Advantage_How_Digital_Leaders_Outperform_their_Peers_in_Every_Industry.pdf)
4. Carlson RH (2011) Biology is technology: the promise, peril, and new business of engineering life. Harvard University Press, Cambridge MA
5. Cham J, Costa K (2017) User experience in R&D. In: PharmaTimes. [http://www.pharmatimes.com/web\\_exclusives/user\\_experience\\_in\\_r\\_and\\_d\\_1203790](http://www.pharmatimes.com/web_exclusives/user_experience_in_r_and_d_1203790)
6. (2016) Gilead Sciences announces acquisition of Nimbus Therapeutics' Acetyl-CoA Carboxylase (ACC) program for NASH and other liver diseases. In: BusinessWire. <https://www.businesswire.com/news/home/20160404005324/en/Gilead-Sciences-Announces-Acquisition-Nimbus-Therapeutics%E2%80%99-Acetyl-CoA>



7. (2011) Hackers, some helpful, penetrate big-name pharma and healthcare. In: FierceBiotech. <https://www.fiercebiotech.com/it/hackers-some-helpful-penetrate-big-name-pharma-and-healthcare>
8. Kiedrowski N (2015) The biotechnology sector continues to defy markets with secular growth. In: Ino.com. <http://www.ino.com/blog/2015/07/the-biotechnology-sector-continues-to-defy-markets-with-secular-growth>
9. Lee T (2014) Hackers break into networks of 3 big medical device makers. In: San Fr. Chron. <http://www.sfgate.com/news/article/Hackers-break-into-networks-of-3-big-medical-5217780.php>
10. Lightbown S (2017) The rise in VC deal sizes since 2012. In: PitchBook. <https://pitchbook.com/news/articles/the-rise-in-vc-deal-sizes-since-2012>
11. M&A deals may pick up for BioPharma in 2017 (2017) In: Morgan Stanley res. <https://www.morganstanley.com/ideas/biopharma-mergers-acquisitions-to-increase>
12. McCandless D, Evans T, Quick M et al (2017) World's biggest data breaches. <http://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/>
13. Meany L (2012) How Nike digitized the #brand experience. <http://www.totalcustomer.org/2012/05/10/nike-digitized-brand-experience/>
14. Mickle T (2017) Oracle pitches security of the cloud to software clients. In: Wall Str. J. <https://www.wsj.com/articles/oracle-pitches-security-of-the-cloud-to-software-clients-1508355668>
15. Shaban H, Nakashima E (2017) Pharmaceutical giant rocked by ransomware attack. In: Washington post. <https://www.washingtonpost.com/news/the-switch/wp/2017/06/27/pharmaceutical-giant-rocked-by-ransomware-attack>
16. Shore R (2015) Massive cancer-drug deal one of UBC's biggest to date. In: Vancouver sun. <http://www.vancouversun.com/health/massive+cancer+drug+deal+biggest+date/11590570/story.html>
17. Symantec Security Response (2015) Butterfly: corporate spies out for financial gain. [https://www.symantec.com/content/en/us/enterprise/media/security\\_response/whitepapers/butterfly-corporate-spies-out-for-financial-gain.pdf](https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/butterfly-corporate-spies-out-for-financial-gain.pdf)
18. Thomson D (2017) US IPOs have their best quarter in two years—Q2 2017 Capital Markets Watch. In: PwC's deal. Blog. <http://usblogs.pwc.com/deals/us-ipos-have-their-best-quarter-in-two-years-q2-2017-capital-markets-watch/>
19. Trends in the cost of computing (2015) In: AI impacts. <https://aiimpacts.org/trends-in-the-cost-of-computing/>