Adapting Extreme Programming to Research, Development and Production Environments

Gil Broza
545 Douglas Ave.
Toronto, ON M5M 1H7 Canada
gilbroza@hotmail.com

Abstract. Affinium Pharmaceuticals engages in early-stage pharmaceutical R&D and molecular biology production processes for internal and external programs. This business requires significant informatics support in terms of small- and large-scale software, tool integration and data management. Obtaining suitable software is difficult due to customer diversity, rapidly evolving unique needs, vendor offering and high costs. Adapting the XP approach and practices for this situation, Affinium’s Informatics group has successfully developed in-house software that has kept up with the science. I describe notable accomplishments, and lessons learned along the way. I propose that a small in-house group of domain-aware developers, using a customized version of XP, would achieve better results than external providers, despite limited access to resources. In closing, I suggest that this structure and methodology are generally applicable to dynamic research, development and production environments.

1 Introduction

Affinium Pharmaceuticals is a structure-guided drug company focused on the discovery of novel anti-infective medicines. Its undertakings thus encompass scientific research, development and production. The drug discovery world is navigated using high-throughput processes for protein production, structure determination, synthesis of chemical matter and myriad assays for drug viability. R&D is responsible for rapidly obtaining drug candidates via proprietary automation, workflows and protocols.

Drug discovery is an extremely expensive and time-consuming pursuit. On average, developing a novel drug costs over US$800-million and takes 15-years from initial research to marketed product [5]. To stay ahead and reach the market faster – and make the most of patent protection – companies must constantly innovate. Affinium’s chosen path is that of utilizing molecular structure, through structural biology, computational chemistry, cheminformatics and bioinformatics. These pursuits are well known for the sheer volume of data they produce and manipulate. Sometimes reaching terabytes per month, this data may yield immense benefits when mined.

Affinium’s line of business requires diverse software for lab and data management, maintaining and accessing scientific databases, streamlining workflow, and operating and integrating instruments. Software is purchased from vendors, outsourced to consultants, or built internally. In the next section I discuss the downsides of purchasing commercial software, and several of the difficulties around planning and engineering useful, affordable solutions, whether done internally or outsourced. I then show how the Informatics group at Affinium has adapted XP to address these problems in building our own software or integrating 3rd-party software.
2 The Problem: Obtaining Appropriate Software

The dynamic nature of the industry poses challenges however a company acquires its software. Systems, which cost a lot of time, money and effort to obtain and integrate, often lose out to innovation and the fast pace of scientific discovery and automation advancement. New business initiatives, or changing priorities in a difficult economic landscape, may trigger obsolescence even faster. Gentler changes, such as new business collaborations or discovery programs, often introduce or modify requirements.

There exist several established vendors of informatics software for biotechnology and pharmaceuticals, but companies in these fields need to do more than choose among them. Given the highly diverse nature of activities in drug discovery, few vendors provide solutions that span the entire spectrum. Some needs are so special that no commercial product addresses them effectively. Lastly, vendors may not be quick, cheap or still around to respond to the changing nature of the business (e.g. [4]).

Data and process integration are vital to a company’s scientific advancement regardless of its size or age. Data is produced and processes are developed in highly specialized departments of different disciplines, such as molecular biology, structural biology and computational chemistry. However, comprehensive data analysis and streamlined discovery operations require integration, posing serious challenges for software development. Achieving useful integration, whether between vendors’ products or custom-built products, is time-consuming and expensive. Cost/benefit analyses lead many companies to opt for a mixed strategy.

A significant downside to eschewing vendors is that a desired system may take many months to enter production. Meanwhile, a lot of money is spent (and a lot more if the project is outsourced), and the requirements mutate. Unformatted or unstructured ‘legacy data’ piles up, which is difficult to use and later import. The scientist customers are often content getting by on their own, for instance by downloading free tools, writing Perl scripts and using Microsoft Excel. Thus, spending money to build software or database systems requires serious justification.

Developer skill-sets are another concern for companies building their own software. The subject matter is difficult to master due to its scientific breadth and specificity, so experts in both software engineering and biology or chemistry are few. Yet, successful developers must have a fairly good exposure to it in order to be conversant with their customers. In my experience, “regular” software developers often find that little of the subject matter strikes close to home for them, unlike financial or telecom systems, for instance. On the other hand, biologists and chemists with programming training often do not have enough experience in building large-scale systems.

3 Our Solution

The Informatics group at Affinium provides internal software development, bioinformatics and cheminformatics services. Unlike the biotech / pharma industry’s tendency to outsource substantially [8], Affinium keeps a small contingent of full-time employees. The benefits have been lower costs and higher quality, greater user satisfaction with the group’s products and availability, and overall better alignment with the company’s goals. In line with industry practice [8], the group comprises 5 to 10 percent of the company’s workforce.