

PART I

GETTING PEOPLE TO COLLABORATE

COPYRIGHTED MATERIAL

1

NEED FOR COLLABORATIVE TECHNOLOGIES IN DRUG DISCOVERY

CHRIS L. WALLER, RAMESH V. DURVASULA, AND NICK LYNCH

- 1.1 Introduction 4
 - 1.1.1 Brief History of Pharmaceutical Industry 4
 - 1.1.2 Brief History of Biotechnology 5
 - 1.1.3 Brief History of Government-Funded Academic Drug Discovery 6
- 1.2 Setting The Stage for Collaborations 7
 - 1.2.1 Current Business, Technical, and Scientific Landscape 7
 - 1.2.2 Externalization of Research: Collaboration with Partners 8
- 1.3 Overview of Value of Precompetitive Alliances in Other Industries 11
 - 1.3.1 Overview of Existing Precompetitive Alliances 12
 - 1.3.2 Pistoia Alliance: Construct for Precompetitive Collaborations 12
 - 1.3.3 How Does Pistoia Plan to Differentiate Itself? 13
 - 1.3.4 Overview of Current Pistoia Projects 14
 - 1.3.4.1 SESL—Semantic Enrichment of Scientific Literature 14
 - 1.3.4.2 Sequence Services 14
 - 1.3.4.3 ELN Query Services 15
- 1.4 Conclusion 17
 - References 17

1.1 INTRODUCTION

From its accidental beginnings in Alexander Fleming's laboratory, pharmaceutical drug discovery and development has emerged as a multi-billion-dollar industry that has revolutionized practically all aspects of human (and animal) life as we know it. Over the past 100 years, serendipitous discovery has been replaced by a structured process that in its current state is highly structured, automated, and regulated. It is also expensive and lengthy and suffers from a 99% failure rate. Industry averages suggest that the cost to bring a new drug to the market under this so-called blockbuster paradigm is in the neighborhood of \$1.5–2.0 billion and takes nearly 16 years (Fig. 1.1) [1].

1.1.1 Brief History of Pharmaceutical Industry

The origins of the pharmaceutical industry can be traced back to the 1800s and the dye industry in Switzerland. From the dye industry, specialty chemistry companies emerged with Ciba, Geigy, and Sandoz in Switzerland along with Bayer and Hoechst in Germany evolving into the first pharmaceutical companies. In the early 1900s, the center of pharmaceutical research and development (R&D) migrated to the United States, specifically New Jersey, with companies such as American Home Products, Johnson & Johnson, Warner Lambert, Merck & Co., Pharmacia-Upjohn, Schering-Plough, BASF, Hoechst, Schering AG, Hoffman LaRoche, and Novartis making it the location of choice for their U.S. operations. The late 1900s saw the emergence of North Carolina as a pharmaceutical industry hot spot with Glaxo-Wellcome making its U.S. headquarters there. Also in the late 1900s, the biotechnology industry emerged

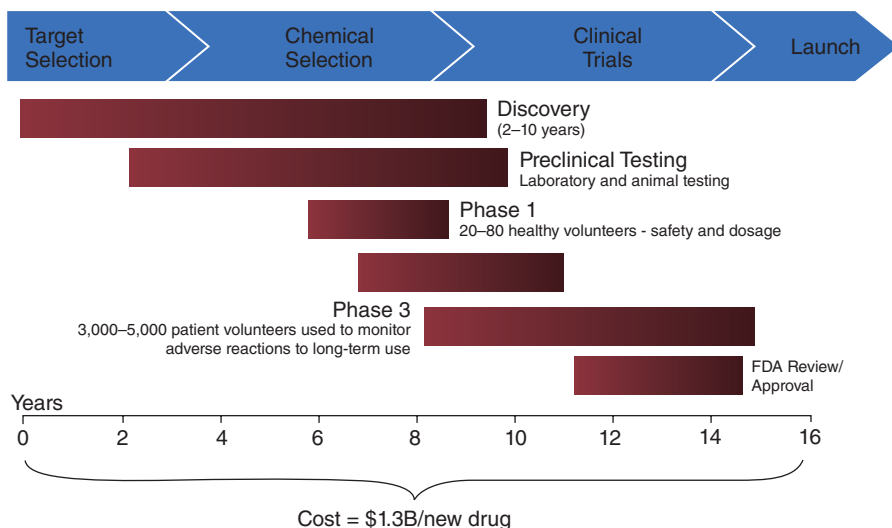


Figure 1.1 Pharmaceutical research and development process.

with companies congregated in the Boston/Cambridge area; the San Francisco Bay Area, San Diego, California; Princeton, New Jersey; Washington, D.C., metro area; as well as Philadelphia. In recent years the economic pressures that forced the pharmaceutical industry to think differently about the sourcing of many operational commodity services has driven a trend toward the emergence of both large pharmaceutical and biotechnology footprints in emerging markets such as Brazil, Russia, India, and China (the traditional BRIC countries) as well as Indonesia [2].

1.1.2 Brief History of Biotechnology

The biotechnology “revolution” began in earnest in 1976 with the founding of Genentech. Inspired by similar movements over the past century in the semiconductor, computer, and advanced materials business, a business model was adopted that would see science evolve from being a tool for the creation of new products and services to being the business itself. Science would move from being “outside” of the business to being the actual business. Genentech was founded as the first of a number of private firms that would monetize the basic research process. Herbert Boyer, an academician, and Robert Swanson, a venture capitalist, invested \$500 each into a new business venture that would seek practical uses for the engineered proteins being developed in Boyer’s laboratory [3]. Genentech remains one of the largest and most successful of the biotech companies, posting revenues in 2008 in excess of \$10 billion, and is now wholly owned by Roche. The Genentech business model continues to be cloned as academicians seek venture capital to advance their ideas and blend science and business.

Despite the business success seen by some of the biotechnology companies, the vast majority of the entrants into this field failed. The business environment imagined (and required) by this new sector was one in which pharmaceutical (R&D) activities were organized through a web of collaborative agreements between the traditional large pharmaceutical and newer biotechnology companies. This collaborative network was envisioned to dramatically alter the industry and transform human health through improved products and services. In reality, while the biotechnology sector has seen exponential growth in revenues over the past 25 years, operational income has been flat or negative, and there has been no discernable difference in research and development productivity as measured by new drug launches. However, the biotechnology sector has contributed to the diversity of treatments in the world’s medicine chest. In 2008, 31 new medicines were launched, 10 of biologics (non-small-molecule) origin, the preferred modality of the biotechnology sector [4].

The promise of transformation of the health care industry brought about by the emergence of “science business” biotechnology companies has failed to materialize due to fundamental differences between the pharmaceutical (R&D) business and the organizational models indiscriminately borrowed

from the semiconductor industry. Science-based businesses face unique challenges not present in these other industries, and the focus on monetization of intellectual property, rather than products or services, has actually been detrimental to the creation of the collaborative network envisioned by the early pioneers of the biotechnology movement. Specifically, this misaligned focus has led to (1) the creation of numerous information silos and barriers to sharing—a key requirement for collaboration, (2) fragmentation of the industry and duplication of noncompetitive activities, and (3) a proliferation of new firms competing for resources from a limited pool [5].

1.1.3 Brief History of Government-Funded Academic Drug Discovery

In 1980, the Bayh-Dole Act was enacted with the intention to stimulate pharmaceutical research into key disease areas by allowing academic institutions as well as individual researchers to benefit directly from commercialization of their government-funded research efforts. Although greatly criticized as a mechanism that promotes science with no direct market relevance [6], government-funded research spending is significant and increasing. Across the National Institutes of Health (NIH), a number of “center grants” have been awarded over the last several years to build out the necessary infrastructure to power an academic revolution. Examples of the types of work being supported are as follows: (1) Burnham was awarded a \$98 million grant to establish one of *four* comprehensive national screening centers as part of the NIH’s, Molecular Libraries Probe Production Centers Network (MLPCN); (2) 83 National Center for Research Resources (NCRR)–funded Centers of Biomedical Research Excellence (COBRE) have been awarded two consecutive, five-year, \$10 million grants; (3) Northwestern is awarded \$11 million to create a Center to Speed Drug Discovery (Northwestern); and (4) a grant from the NIH will help establish the Chicago Tri-Institutional Center for Chemical Methods and Library Development. The NIH will pump \$62 million into more than 20 studies focused on using epigenomics to understand how environmental factors, aging, diet, and stress influence human disease.

In 2008, the National Cancer Institute (NCI) alone funded research efforts in excess of \$12 billion. More recently, the NCI has been funding efforts that would increase the value of academic research through the creation of public–private partnerships to translate knowledge from academia into new drug treatments. To this end, the NCI has established the Chemical Biology Consortium, which is advertised as an integrated network of chemical biologists, molecular oncologists, and chemical screening centers. Current members of the consortium include. The University of North Carolina in Chapel Hill, North Carolina; Burnham Institute for Medical Research in La Jolla, California; Southern Research Institute in Birmingham, Alabama; Emory University in Atlanta; Georgetown University in Washington, D.C.; the University of Minnesota in St. Paul and Minneapolis; the University of Pittsburgh and the University of Pittsburgh Drug Discovery Institute; Vanderbilt University

Medical Center in Nashville, Tennessee; SRI International in Menlo Park, California; and the University of California at San Francisco.

Like the biotechnology revolution of the late 1970s, the current trend in the creation of networks of public and private institutions, if successfully operationalized, could transform the health care industry. It is important to acknowledge the lessons from the biotechnology revolution as discussed above and plan accordingly to avoid the pitfalls. In order to be successful, the academic institutions must strive to establish truly open and standard data exchange mechanisms and coordinate activities effectively across a highly distributed enterprise that must adopt an integrated business process.

1.2 SETTING THE STAGE FOR COLLABORATIONS

A reorientation of our business models to focus on products and services will be required if the collaborative R&D environment is to be effectively realized. An acknowledgment, by the industry as a whole, must be made that we differentiate ourselves in the marketplace not through our intellectual property but rather through the delivery of products and services that attract and retain consumers. The R&D process, in any industry, is timely, expensive, and, except for those rare instances where true discoveries/inventions are being made, commoditizable across the industry in the sector. A clear understanding and declaration of what differentiates one company from the next in the marketplace must be established and adopted. Only then can we begin to pool our limited resources effectively to solve common problems and focus our specific internal resources on the elements of the R&D process that allow us to transform the health care system and succeed in the marketplace as individual companies.

1.2.1 Current Business, Technical, and Scientific Landscape

The business value of an information technology (IT) system is based on the ability of the system to support and enhance the business process. Fundamentally, open standards are intended to provide resilience to withstand the technical volatility within business processes and their associated systems. If a system and the business process were flawlessly stable over many years, then there would be little value in developing and adopting standards. However, within the pharmaceutical industry, volatility and upheaval abound in every phase of R&D. Perhaps the largest source of upheaval within our industry is the volatility of mergers and acquisitions (M&A) among industry peers as well as business partners, commercial suppliers, and clinical research organizations (CROs) (Fig. 1.2). This M&A volatility—coupled with exponential growth in outsourcing—has placed tremendous pressure on R&D processes to change frequently and dramatically. Common pharmaceutical processes like target identification, compound synthesis, in vivo toxicology, biomarker discovery,

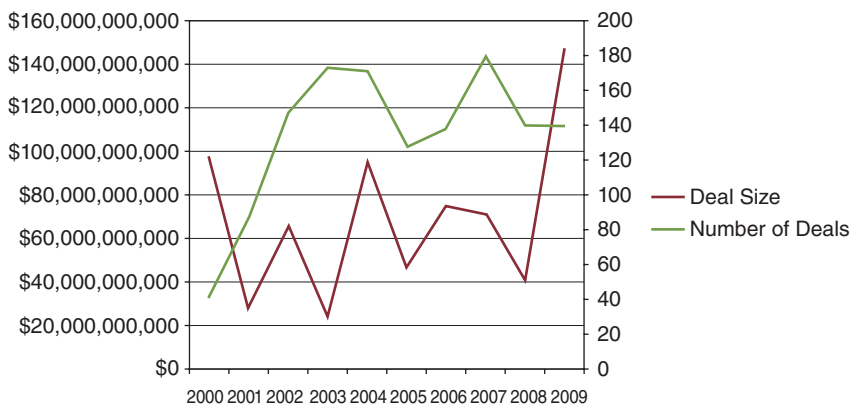


Figure 1.2 Pharmaceutical M&A activity, 2000–2009. (Source: <http://www.marketwatch.com/story/ten-year-data-on-pharmaceutical-mergers-and-acquisitions-from-dealsearchonlinecom-reveals-top-deals-and-key-companies-2010-03-25>. MarketWatch data based on original content from DealSearchOnline.com.)

patent searching, and pharmaceuticals are all experiencing revolutions in their processes. The related systems are thus also reacting to this process volatility. This upheaval in the requirements and specifications of R&D IT systems is causing IT budgets to increase, exactly at the moment when all budgets across R&D are sharply decreasing.

We face an unprecedented era of rising process upheaval and constantly evolving business requirements coupled with a cost-conscious environment where chief information officers (CIOs) and R&D executives are looking to simplify their IT architectures and their cost basis. If this trend continues, informatics systems may become a bottleneck to the productivity of pharmaceutical scientists.

1.2.2 Externalization of Research: Collaboration with Partners

The area of greatest process upheaval is the externalization of research processes and the growing collaborations between life science partners throughout the R&D cycle. Originally CROs had been outsource partners, but currently there are outsourcing partners for every phase of the R&D process, from target identification to chemical synthesis to pharmacokinetic studies to clinical supplies, and so on. With this increased opportunity and necessity for outsourcing, samples are constantly getting shipped to and from pharmaceutical laboratories. Every time a sample changes hands, there is a related data exchange as well. Often, for a pharmaceutical company, several CRO partners will be used for a single research project. Also, the CRO will likely have several pharmaceutical clients. In this emerging net-centric industry model, there is a complex graph of data exchange that must be supported (Fig. 1.3).

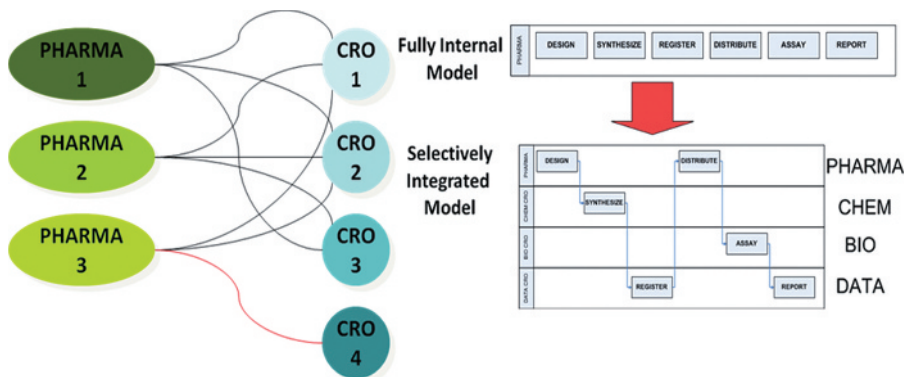


Figure 1.3 Emergence of a selectively integrated drug discovery and development model.

For example, for every pharmaceutical company, there may be two or three chemistry synthesis partners. These partners would likely have their own internal systems for tracking reagents, recording experiments, and registering novel compounds. Since the synthesis is performed on behalf of the pharmaceutical client, a majority of the data from the experiment, from reaction yields to analytical data, must be transmitted to the client along with the synthesized compound in a vial. The challenge is that since the pharmaceutical client has developed mature internal processes, and the synthesis partner has its own internal processes, there is a high likelihood that the processes—and the related IT systems—are different in nature. This leads to the use of different metadata, different vocabularies, and different quality control on the data capture. When an instance of a novel compound is synthesized, the outsource partner may call it a “batch” but the pharmaceutical client may call it a “lot”. Also, some compound registration systems assign a different identifier for different salt forms of the compound. One company may handle this by using a suffix of the compound identifier (<compound identifier>-<salt form>), whereas another company may simply assign a completely different base compound identifier to the different salt form. Both of these are legitimate taxonomies to register and identify compounds and their salt forms. The difficulty comes when one company attempts to export its registration data and transmit that to the other company. Reconciling the differences in the semantics and vocabularies of different compound registration systems can be a tedious, error-prone, and often irreconcilable task. Often this reconciliation involves compound registrars and synthetic chemists (and possibly lawyers) from both parties. If the need to transmit compound registration data between business partners was a unique event, then perhaps a manual reconciliation process would suffice. However, since every pharmaceutical company has several synthesis outsourcing partners, and every synthesis CRO has several pharmaceutical clients, this metadata-conflict and reconciliation process is

repeated over and over throughout the industry. While this problem of data reconciliation and reformatting is time consuming and error prone in the chemical synthesis domain, this problem is often even more exacerbated in the biological domain.

Often pharmaceutical companies will have outsourcing relationships with contract laboratories that perform assays on compounds owned by the client. These assays could be standard assays that are outsourced for cost efficiencies or proprietary assays that are otherwise not available to the pharmaceutical client. As with compound registration systems, the outsource partner that runs the assays will likely have internal protocol registration and biological assay data management systems to capture the data. These systems will be built to suit the needs of the internal processes within the contract laboratory, so that they can properly manage, interpret, and report on their assay results. However, most pharmaceutical companies like to import the assay results into the pharmaceutical company's internal assay data management system. This would enable the pharmaceutical scientists to interpret the outsourced assay data side by side with all of the other data generated on that proprietary compound. With every partner that generates assay data related to a compound, there is an ongoing, complicated effort to properly format and transmit the data such that the scientists in the pharmaceutical company can understand the nature of the assay and accurately interpret the results. Too often, many days are wasted merely explaining differences between internal and external assay results. Especially with high-throughput or high-content biological assays, there are a significant number of attributes of the experimental design that are important to account for in the data interpretation. For example, which cell line was used? Was it a single-point assay or a dose-response? What was the detection mechanism; fluorescence, phosphorescence, and so on? Furthermore, there are many cases where the proprietary assay platform generates data that have a unique structure.

Perhaps the assay is a high-throughput, low-resolution format, in which case the raw numeric output must be binned into low-medium-high categories and only the binned values are reported to the client, yet the client has stringent data quality, numbers-only rules to which the contract laboratory cannot adhere. Perhaps the assay has a cutoff at a reading threshold, causing the result to be reported as a range instead of an explicit number. Perhaps there is a nonlinear response that requires special curve-fitting software to calculate the half maximal inhibitory concentration (IC_{50}) value. There are many nuances and subtleties to biological assay data, and a large amount of metadata is required to properly describe the experimental method. This must be understood by the scientist who is using that assay data to make design or synthesis decisions for the next molecule. As such, it is important for the contract laboratory to deliver the full experimental description of its data and for the pharmaceutical customer to ingest and report all of that description to its scientists. Again, as with compound synthesis, if this assay data generation was done with a single partner, then a manual process with significant interactions between

business partners would be appropriate. However, pharmaceutical companies often send their compounds to many laboratories to be tested in numerous assays, and all of that data must be imported into the assay database of the client, and the data must be interpreted by chemists and biologists who are not the operators of those assays. The further downstream the assay if the assay was an *in vivo* assay, as opposed to an *in vitro* assay—the more complicated the experimental design, and thus the harder it is for scientists to interpret the data without being proximal to the biologist who performed the assay.

Both the chemistry and biology examples above highlight the cost and complexity of exchanging data between business partners, and the activities of data exchange and data harmonization are not value-added work for finding drugs. These data tasks are a cost of doing business in life sciences, and as such the industry is looking for ways to reduce these costs without impacting the science. In fact, it could be argued that resources poured into the data activities are actually *diverting* funds away from doing science. So, reducing these costs will actually free up resources to do more science. The challenge of reducing these data-curation costs is that no single entity, neither a pharmaceutical company nor a contract laboratory nor a biotech, can accomplish what is needed to be done, namely to harmonize across the industry. Point-to-point optimizations of data exchange are helpful but only marginally cost effective. For a paradigm shift to occur that would dramatically improve the efficiency of external science, the industry must come together to agree on common methods of exchanging data, delivering services, defining entities, and so on. Thus, a precompetitive collaboration among informatics groups is a natural evolution in our industry. This evolution has already occurred in numerous other industries, from apartments [7] to banking [8] to retail [9].

The nature of every industrywide data standardization effort revolves around defining the terminology, semantics, metadata, entity attributes, and services or functions of the data exchanged between business partners. These definitions and attributes are collaboratively defined by IT or informatics peers who together determine how to harmonize data between disparate systems and processes.

1.3 OVERVIEW OF VALUE OF PRECOMPETITIVE ALLIANCES IN OTHER INDUSTRIES

Other industries have realized the need for precompetitive alliances for some time and have established them over the last two decades. This drive for collaborative alliances has been driven by the same pressures that the life science industry faces today, that of increased pressures on efficiency and the need to divert funding to innovative activities rather than to commodity services. The maturity of the business model for these other industries (telecoms, insurance, automotive, and aerospace) has meant that they have existed prior to work within the early stages of life science and informatics. These other industries

realized early on that each company existed as part of an extended ecosystem that relied on the ability to do business with other partners and competitors and hence where the need for interoperable processes and information flows were critical to their mutual success.

1.3.1 Overview of Existing Precompetitive Alliances

Without going into details on all the other industries, some have direct parallels with discovery life science from both other life science areas and financial services. The financial services industry created the VISA processing standards and in creating this concept has led to an explosion in the ways that credit cards are used and their ease of interoperability. Other examples of open approaches include the insurance industry (Polaris) to support data exchange between insurance brokers and the insurance companies offering the policies. In the clinical development workflow of development pharmaceuticals the need to work with multiple partners as part of the delivery of clinical trials and the later delivery of health care services to patients has provided the environment for groups such as the Clinical Data Interchange Standards Consortium (CDISC: www.cdisc.org) and Health Level 7 (www.hl7.org) to be founded and evolve over several years. The drivers here were a need for interoperable standards for information delivery and data markup to support effective and clear communication for submission of clinical trials data and the later management of health care information.

The way these companies do business has changed as the global economy has evolved, but delivering critical information to scientists continues to be the key part of the R&D informatics groups within these pharmaceutical and agrochemical companies and support organizations. There are various ways that the development of software and delivery of information to scientists can be improved through collaboration and open standards. There is evidence from other global businesses where strong open standards have benefited a whole industry sector and delivered improved innovation in the face of cost pressures.

1.3.2 Pistoia Alliance: Construct for Precompetitive Collaborations

There has been a history of organizations working together to promote common standards in the early-stage life science industry over the last decade both as new groups established specifically for life science [Interoperable Information Infrastructures Consortium (I3C: www.i3c.org), Society for Bimolecular Sciences (SBS: www.sbs.org), BioIT Alliance (www.bioitalliance.org)] and those attached to larger groups but wishing to explore and adapt into life science [Object Management Group (OMG: www.omg.org), World Wide Web Consortium (W3C: www.w3c.org)]. The success rate has been variable over the years with various initiatives coming and going and others building a portfolio of activities and evolving. Much of the thinking of setting up the Pistoia

Alliance (www.pistoiaalliance.org) has tried to take the learning from these other groups and understand how they were able to deliver collaborative value.

1.3.3 How Does Pistoia Plan to Differentiate Itself?

There are various factors that we believe make the Pistoia Alliance work slightly differently, including a changing economic environment that is forcing more collaboration and improvements in software design that focus on software services which allow a high level of abstraction and hence more opportunity for cross-company integration. The high-level business processes executed within this sector are very similar between different organizations, and the further appreciation that there is considerable overlap and commonality in the processes executed within the sector has made groups question what is competitive advantage and what are supporting assets that could share some common design (Fig. 1.4).

A key element for the establishment of the Pistoia Alliance was ensuring that the life science business needs were the driving force for the development of common standards and approaches in the group rather than simply a technology/solutions focused view. Hence the projects that have evolved in the first build of the Pistoia Alliance program are intended to show these drivers from developing service requirements (sequence services) and an open

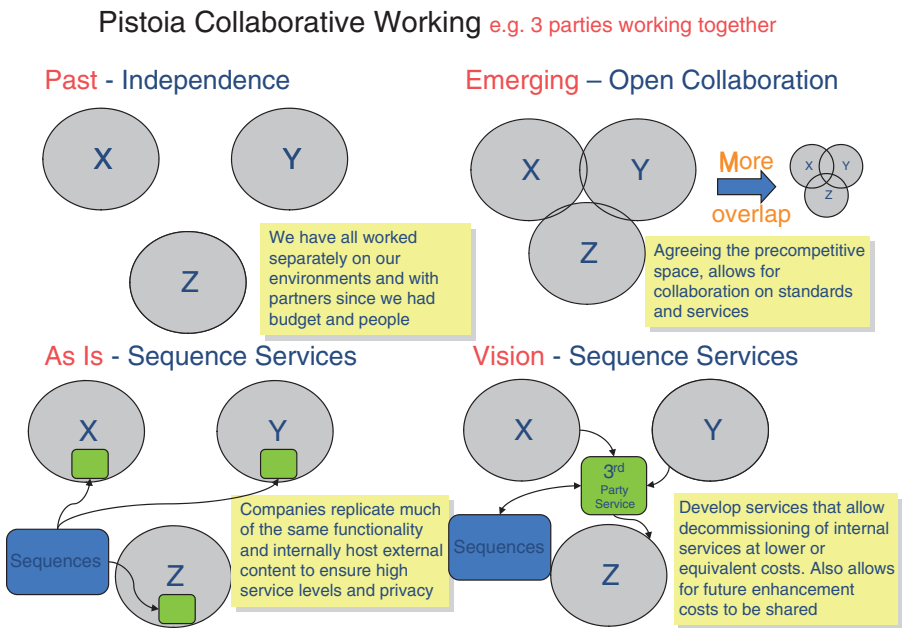


Figure 1.4 Pistoia Alliance collaborative working model.

framework based on existing standards (SESL). The key intention of the Pistoia Alliance was to move beyond standards in their adoption as service requirements and into influencing future business models and be a potential for change in the delivery of information and services in the life science industry. The next-generation business model would ideally shift from products (software programs or databases that need to be installed and maintained) to services (accessing data on Web-based platforms or hosted off-site), eventually maturing to “software as a service,” known as SaaS, which would be deployed over the Internet. Standard interfaces, such as those used by Web browsers, would make it easier to simplify IT architectures across the industry, and centralized services would deliver economies in scale and scope. Among the major benefits would be reductions in cost and maintenance as information silos inside company networks are turned off in favor of fewer, more versatile tools. The Alliance has a broad membership because such extensive changes in the business model affect all parts of the supply chain, from life science back to software providers and content providers.

We want to have all parties [suppliers, academics, nongovernmental organizations (NGOs), pharma, and life science companies] actively involved in the Alliance’s initiatives, as the intent is to deliver practical pilots and prototypes that demonstrate the collaborative activity. The Pistoia Alliance differentiates itself from groups both past and present through its attempts to embrace and extend the standards and services of these companion groups in technology offerings driven by clear business needs. We wish to adopt existing standards where we can rather than create new ones and also collaborate with existing groups to bring fresh ideas into the value chain. We list a selection of our current portfolio that highlights our current foci and also the wider impact on the information delivery models.

1.3.4 Overview of Current Pistoia Projects

1.3.4.1 *SESL—Semantic Enrichment of Scientific Literature* The Pistoia Alliance project on biomedical knowledge brokering standards (SESL) is developing a pilot to showcase its key approaches, and its aim is to demonstrate the feasibility of an open knowledge brokering framework which will reduce the costs of integration of disparate data types from several sources. The pilot is focused on the extraction of assertions for type II diabetes mellitus (T2DM) from both the scientific literature, supplied by participating publishers, and structured data resources managed by EMBL-EBI (the European Bioinformatics Institute). The pilot [expected to include an (resource description framework (RDF) triple store)] will be published and a prototype demonstrator will be made publicly available to show feasibility (Fig. 1.5).

1.3.4.2 *Sequence Services* Most major pharmaceutical companies currently host a large number of sequence data and analysis tools within their firewalls. While the genome was still being sequenced, and during the race to patent

Biomedical Knowledge Brokering

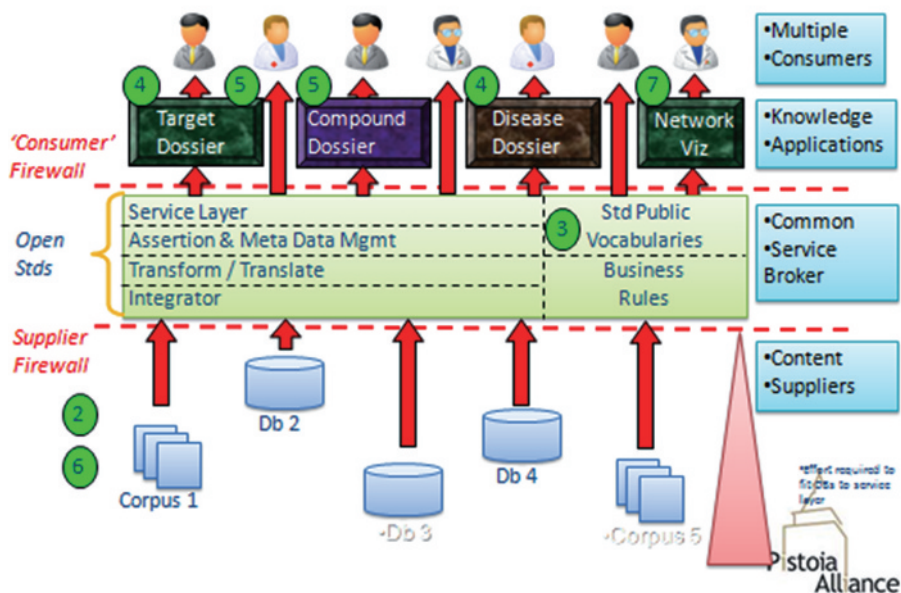


Figure 1.5 Schematic Architecture for SESL project.

genes, these services offered a competitive advantage, and consequently each company built and maintained vast internal systems that both took external public data and merged it with internal private data. However, in the past five years the public domain has caught up (and in many cases surpassed) the expensive, heavily customized commercial and proprietary solutions used by industry.

As a drive to cut costs, encourage standards, and provide simplification, the Pistoia Alliance is commissioning a pilot set of secure hosted sequence services based on the functional and nonfunctional requirements of its members. These services will provide access to public, private, and commercial data and tools that will enable scientists to search, store, and analyze all their sequence-based data in a single Web interface. Additionally data will be searched and accessed via Web services to allow sophisticated users to flexibly retrieve or pipeline data (Fig. 1.6).

1.3.4.3 ELN Query Services The adoption of an electronic laboratory notebook (ELN) within an organization is as much a business change process as it is a technology project, and so the ELNs have traditionally had to focus on the role of the experimental scientist entering new information and ensuring this process is managed and efficient. In areas where ELNs have been used for a few years, such as supporting chemistry synthesis (medicinal chemistry,

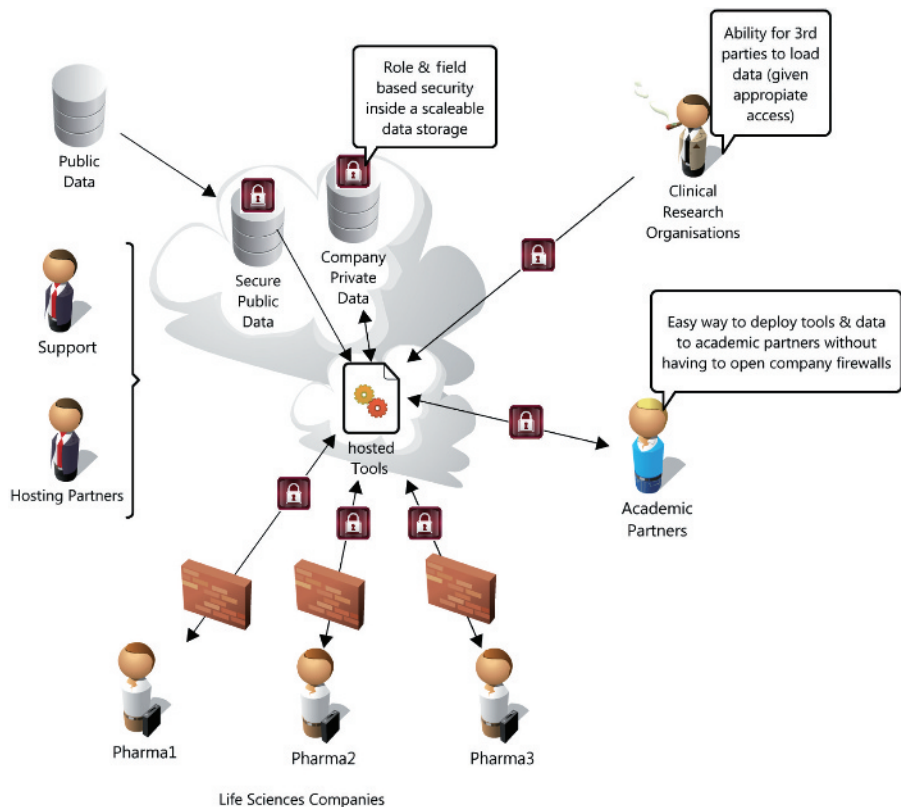


Figure 1.6 Conceptual view for the sequence service project.

process chemistry, operations, and manufacturing), there is a growing demand for enhanced exploitation of the data held within an ELN and the future linking of that data with relevant data held within an organization or further afield. The requirements for knowledge management have grown considerably in the last few years, and this increases the need to query the ELN to extract the high-value information and to build assertions with other data from within an organization or outside (Fig. 1.7).

As the number of ELN installations grows, this requirement becomes more challenging, particularly given the diversity of such ELN implementations (developed commercially, in-house, blended, or as open-source systems). In many companies already a mixture of ELNs have been deployed, either through conscious choice or as a result of mergers and acquisitions. Another key factor is the trend for more business process outsourcing, resulting in the need to be able to work with a CRO partner and share aspects of an ELN knowledge base. So the problem the industry faces is twofold: (1) the need for

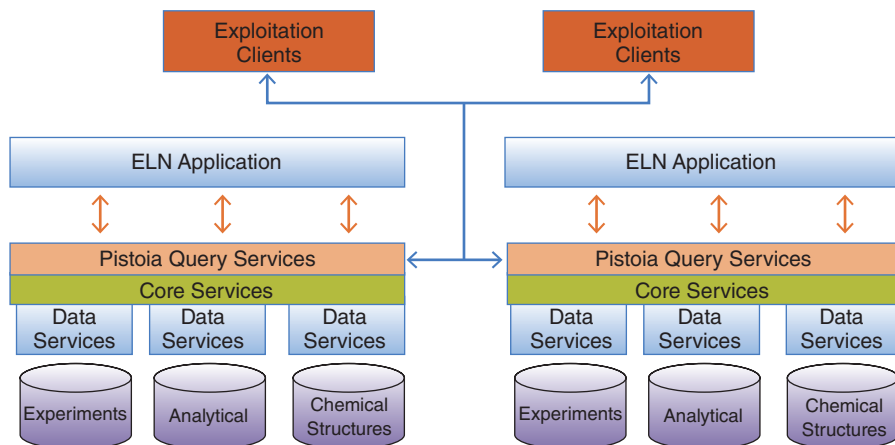


Figure 1.7 Conceptual vision for ELN project.

better exploitation of ELN data and (2) the need to build different ELN implementations using different domain models and designs.

1.4 CONCLUSION

A precompetitive collaboration, the Pistoia Alliance, has been established to provide the foundation of data standards, ontologies, and associated Web services to enable pharmaceutical discovery workflow through common business terms, relationships, and processes. The initial focus has been on chemistry, biological screening, and sample logistics. All pharma companies and software vendors are challenged by the technical interconversion, collation, and interpretation of drug/agrochemical discovery data, and as such, there is a vast amount of duplication, conversion, and testing that could be reduced if a common foundation of data standards, ontologies, and Web services could be promoted and ideally agreed upon within a nonproprietary and noncompetitive framework. This would allow interoperability between a traditionally diverse set of technologies to benefit the health care sector.

REFERENCES

1. *Outlook 2010*. Boston, MA: Tufts University, 2010.
2. Research and development in the pharmaceutical industry. Washington, DC: U.S. Congressional Budget Office, 2006.
3. WGBH. Herbert Boyer. *They Made America*. Boston: PBS, 2004.

4. PhRMA Annual Report 2009. Pharmaceutical Research and Manufacturers of America, 2009.
5. Pisano GP. *Science Business: The Promise, the Reality, and the Future of Biotech*. Boston, MA: Harvard Business Press, 2006.
6. Loewenberg S. The Bayh-Dole Act: A model for promoting research translation? *Mol Oncol* 2009;3:91–92.
7. Multifamily Information and Transactions Standard. Available: <http://www.mitsproject.com>.
8. Transactions Workflow Innovation Standards Team. Available: <http://www.twiststandards.org>.
9. The Association for Retail Technology Standards. Available: <http://www.nrf-arts.org>.