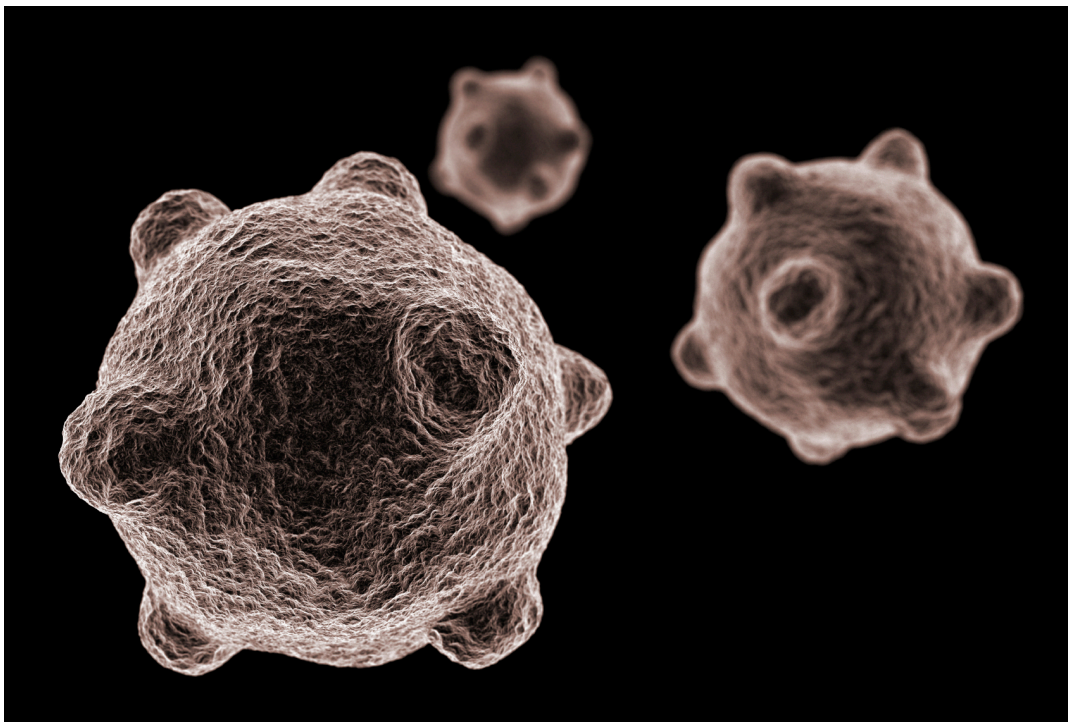


THERAPY DEVELOPMENT IN A NETWORKED WORLD

INTRODUCTION AND OVERVIEW

MARTY TENENBAUM AND JOHN WILBANKS

May 2008



## Drug Discovery: Time for a Change

The pharmaceutical industry is at a crossroads. Despite revolutionary advances in molecular biology that have made genetic decoding routine, the time from gene to cure still stands at 17 years. High-throughput screening methods allow us to test the efficacy of millions of compounds against a molecular target in a single week; but the odds of one of those compounds making it through the development pipeline and becoming a drug are less than 1/1,000,000. A well-funded group starting today, using the traditional model of drug development, has a very slim chance at getting a drug to market by 2025.

The time has come to change the way we cure disease. We are no longer asking whether *a* gene or *a* molecule is critical to a particular biological process; rather, we are discovering whole networks of molecular and cellular interactions that contribute to disease. And soon, we will have such information about *individuals*, rather than the population as a whole. Biomedical knowledge is exploding, and yet the system to capture that knowledge and translate it into saving human lives still relies on an antiquated and risky strategy of focusing the vast resources of a few pharmaceutical companies on just a handful of disease targets.

The current path to drug discovery, furthermore, relies on an inherited distribution of risks and rewards that actually impedes progress. In academia, pre-clinical research is rarely rewarded, and few labs possess the resources needed for industrial-scale development. Translational research fares no better in industry, where the inherent risks are incompatible with the way life sciences businesses are organized, managed and financed. Hundreds of biotechnology companies and dozens of large pharmaceutical companies operate through loose milestone-based alliances instead of long term relationships - an ecosystem that does not foster integration and shared learning.

The current path to drug discovery also perpetuates old traditions of information and intellectual property control. This deeply set inability to capture collective learning dooms everyone to revisit infinitely many blind alleys. The currency of scientific publication encourages individual scientists to hoard rather than share data that they will never have the time or resources to exhaustively mine. And, the wealth of “negative” information gleaned from clinical trial data is mostly lost to the need for companies to safeguard their commercial investments. Although computational and systems biology, aided by Moore’s law, make it feasible to systematically search the vast space of targets, leads, and interactions, this potential is limited in practice by lack of access to data, compound libraries, specimens, and shared

services essential for economies of scale. As a result, many biological promising leads, and the knowledge surrounding them, are ultimately discarded.

Thankfully, we have a rare moment in time where we can change the entire system in one motion by establishing a collaborative ecosystem of knowledge and research services that can be rapidly assembled to develop new therapies with unprecedented efficiencies and economies of scale. We can create the same radical increase in efficiency for scientific research that commerce saw in the 1990s, as secure Internet transactions transformed many vertically integrated industries into horizontally integrated ecosystems of service providers and consumers. The explosion of contract vendors in biotechnology, covering the spectrum from gene to protein to drug discovery, development and trials, is one factor. The emergence of the Semantic Web for science is part of the story, as is the existence proof that common use licensing can create explosive value in software and culture. And the power of the network to bring these elements together into a unified system, a *Health Commons*, is the final piece of the puzzle.

### **The Health Commons Vision**

Imagine a virtual marketplace or ecosystem where participants share data, knowledge, materials and services to accelerate research. The components might include databases on the results of chemical assays, toxicity screens, and clinical trials; libraries of drugs and chemical compounds; repositories of biological materials (tissue samples, cell lines, molecules), computational models predicting drug efficacies or side effects, and contract services for high-throughput genomics and proteomics, combinatorial drug screening, animal testing, biostatistics, and more. The resources offered through the Commons might not necessarily be free, though many could be. However, all would be available under standard pre-negotiated terms and conditions and with standardized data formats that eliminate the debilitating delays, legal wrangling and technical incompatibilities that frustrate scientific collaboration today.

We envision a Commons where a researcher will be able to order everything needed to replicate a published experiment as easily as ordering DVDs from Amazon. A Commons where one can create a workflow to exploit replicated results on an industrial scale – searching the world's biological repositories for relevant materials; routing them to the best labs for molecular profiling; forwarding the data to a team of bioinformaticians for collaborative analysis of potential drug targets; and finally hiring top service providers to run drug screens against those targets; with everything – knowledge, data, and materials – moving smoothly from one

provider to the next, monitored and tracked with Fed-Ex precision; where the workflow scripts themselves can become part of the Commons, for others to reuse and improve.

Health Commons' marketplace will slash the time, cost, and risk of developing treatments for diseases. Individual researchers, institutions, and companies will be able to publish information about their expertise and resources so that others in the community can readily discover and use them. Core competencies, from clinical trial design to molecular profiling, will be packaged as turnkey services and made available over the Net. The Commons will serve as the public-domain, non-profit hub, with third-parties providing value added services that facilitate information access, communication, and collaboration.

By accelerating the pharmaceutical industry's transition to a horizontally integrated ecosystem of service providers, Health Commons will usher in a new age of collaborative drug and therapy development, reducing costs and risks through economies of scale and cross-learning. It will focus resources on key opportunities, disseminating new results proactively, and minimizing unnecessary replication. It will encourage outsourcing to centers of excellence, with attendant economies of scale. It will broaden access to unique experimental resources. It will promote a culture of continuous process improvement and learning. It will enable a small group of individuals with modest funding to replicate the existing functions of a venture-backed biotech company. This may sound astounding, but it's no more astounding than the idea that a small bunch of individuals might start an internet business to compete with the biggest names in bookselling sounded in 1992.

And like the Web, it will unleash an explosion of innovation, by encouraging everyone to build on each others' services, creating a myriad of new ones.

### **What is Health Commons?**

Health Commons is a coalition of parties interested in changing the way basic science is translated into the understanding and improvement of human health. Coalition members agree to share data, knowledge, and services under standardized terms and conditions by committing to a set of common technologies, digital information standards, research materials, contracts, workflows, and software. These commitments ensure that knowledge, data, materials and tools can move seamlessly from partner to partner across the entire drug discovery chain. They enable participants to offer standardized services, ranging from simple molecular assays to

complex drug synthesis solutions, that others can discover in directories and integrate into their own processes to expedite development -- or assemble like LEGO blocks to create new services.

Health Commons builds on the existing public domain of content and tools, supporting the enormous resource of the Entrez-NCBI system, as well as other public-private efforts like the HapMap, Allen Brain Atlas, Alliance for Cellular Signaling, Biomarker Consortium, Neurocommons, and emergent efforts in chemistry and toxicity. Simply putting digital resources on the Web, however, is not enough. Health Commons will implement community standards in data descriptions, so that information generated in one part of the Commons can be re-used elsewhere without costly negotiations over data formats and scientific meaning. Providing such standards, Health Commons improves and extends the public domain by integrating hundreds of public databases into a single framework, and lays the foundation for dramatic increases in the number of openly available content and tools, and their utility for collaborative investigations. This “open source data integration” lowers the costs of using the public domain at the same time that it increases the utility of public information.

The Commons also requires social and legal infrastructure, such as standardized legal agreements and pre-negotiated licenses that streamline the sharing of materials, data, knowledge, and IP among all Health Commons participants. For example, Commons-enabled compound libraries might be made available under a “some rights reserved” research model with pre-negotiated access agreements. Standard contracts also improve the efficacy of tissue banks and other similar resources, making previously difficult-to-acquire research materials available for open order. Legal standards include “materials transfer agreements” that allow expansive research rights, but no commercialization rights, to all parties, as well as agreements subject to post hoc “credit assignment” by an objective third party if and only if a collaboration is successful. The Health Commons legal framework is based on a simple but profound principle: time-consuming and costly negotiations over IP ownership, royalties and similar issues should be deferred until something proves worthy of negotiation.

The final element of Health Commons is a Web-based collaborative research platform that enables participants to discover, access, integrate, and share the resources they need. This platform, which combines elements of a data and knowledge sharing hub, search engine, workflow engine, ecommerce site, and collaborative workspace, serves as the public portal to Health Commons. It enables research teams to share data, knowledge, clinical resources and services; automate and track workflows, collaborate on data analysis, and coordinate research priorities and resources.

The platform is open and standards-based so it can be integrated with popular research tools and environments. Equally important, users can build their own private Commons (analogous to an Intranet), allowing individuals and companies to leverage the power of the Health Commons, either as a hosted service or behind their own firewalls. Individuals can create private workspaces in which to analyze and interpret their data using publicly available resources.

The uniform platform architecture underlying all Health Commons communities allows 1-click “publishing” (i.e., sharing) of private information and resources -- initially to one’s lab, then to collaborators, and ultimately to the greater research community. Such sharing can be actively encouraged by wiring it into workflows as a default that is triggered by events such as the acceptance of a paper, or the failure or abandonment of an experiment. This seamless transition, from private information through stages of closed collaborations to public information, has two important implications: it encourages early adoption by individual researchers, to accelerate their own work; and it enables Health Commons to grow organically as a family of independent but compatible research communities, each contributing data and services to the public Commons, and collectively generating powerful network effects.

### **Health Commons transforms drug discovery.**

We know that existing monolithic discovery models are broken. Current models are not just failing to discover drugs, they're failing to contribute meaningful research into a place where it can help with the next experiment. The answer isn't more machines and more data. It's nothing less than reworking the entire business model for therapy development – processes, economics and culture -- into a model that exploits the power of collaboration and networks to radically reduce the time, costs and risks of therapy development.

To understand just how broken things are and how a Health Commons can help, consider the case of rare, orphan and neglected tropical diseases. Well over 5000 such diseases are simply uneconomical for traditional pharmaceutical and biotechnology companies to address. While academic researchers often work on such diseases, the cost, scale, and complexity of drug development is beyond the reach and expertise of academic labs. So many promising targets and leads wallow in technology licensing offices, never to see the light day. Health Commons makes it cost effective for small groups of researchers to conduct industrial scale R&D on rare diseases by exploiting the economies of scale afforded by an ecosystem of shared knowledge

and services. In the age of genomic sub-typing and personalized medicine, where every disease is potentially an orphan disease, the implications are far-reaching.

Occasionally an opportunity appears so promising that the researchers spin out a biotechnology company. But the costs are high and the odds of success dismal. Integrating resources from the Commons into a virtual biotechnology collaboration offers an attractive way of bridging the “valley of death” from lab to clinic by significantly reducing the time and cost of validating and developing promising opportunities. Virtual Biotechs are much faster and cheaper to set up, since they can reuse workflow and expertise of predecessors. Pooling IP under a standard Health Commons agreement focuses everyone’s attention on finding cures, not negotiating contracts. And if development hits a dead end, a Virtual Biotech can be taken apart just as quickly as it was set up – or placed in “hibernation” to await the development of new science without burning cash.

In fact, many research foundations have recognized the need for just such a model. Some 2500 foundations invest hundreds of billions annually seeking cures to hundreds of diseases, yet, there is little coordination, resource sharing or cross learning among initiatives and no effective processes for moving research into the clinic. Targeted research foundations like the Myelin Repair Foundation, which develop detailed research agendas across cross-disciplinary scientific collaborators, have emerged as alternatives to the massive disease foundations. Health Commons is precisely the platform these foundations need to crystallize their stable of researchers into an effective collaborative community for translational research. Collectively, foundations will benefit from unprecedented opportunities for economies of scale and cross learning across diseases.

But the benefits of Health Commons extend far beyond the non-profit world. Pharmaceutical companies currently prosper by maximizing the value of leads and targets on which they hold commercial rights or patents, not by curing particular diseases. Promising targets and leads for a disease that don’t conform to their business needs are usually ignored. Every pharmaceutical company sits on a wealth of promising targets and leads that they won’t develop themselves. There are also strong competitive incentives against sharing data (even failed results) or drug libraries with competitors, resulting in needless repetitions of failure. That’s not just bad business – it’s tragic. Everyone is behaving rationally, but the incentives are perversely aligned against finding timely cures.

Health Commons can raise pharmaceutical productivity and profits by re-aligning incentives to encourage such collaboration. While a large part of a company’s intellectual property will

necessarily remain proprietary, what is to stop us from routinely sharing databases and predictive models of efficacy and toxicity of chemical entities, as well as data on failed drugs and trials? Sharing toxicity and trials data would improve the ability to “fail fast,” and more fundamentally improve our collective understanding of human toxicity and drug response. “Failed” clinical trials are no longer failures, but the seedbeds for new understandings of how chemicals interact with cells. Several successful examples of such cooperation have already emerged, with hundreds of millions of dollars invested in mapping genetic variation and understanding cell signaling and new models like the Diabetes Genetics Initiative (DGI), a pioneering public-private collaboration of The Broad Institute, Lund University, and Novartis.

Beyond improving our scientific understanding, there are more immediate and compelling reasons to share data from clinical trials – far and away the largest sink of time and money in drug development. Increasingly, failed drugs are finding success as targeted therapies for a selected subclass of the disease for which they were developed, or for a completely new indication not originally considered. These salvage opportunities result from a re-analysis or meta-analysis of the trial data, often by independent investigators, to identify subclasses of responders and understand the unintended consequences of “side effects”. Unfortunately, it is currently very difficult for researchers to obtain clinical trial data. Privacy concerns on behalf of patients, liability concerns on behalf of companies, and more make clinical data a rarely shared commodity.

Health Commons can solve this problem with the help of trusted escrow agents from the government and banking sectors. De-identified data can be combed by academic partners under standard, pre-negotiated terms to find and publish sets of biomarkers associated with disease and toxicity; a standard set of biomarkers will help companies much more finely estimate the odds of success in-clinic, decreasing the number of times a likely failure survives the risk management gauntlet. The Health Commons will dramatically increase the amount of meaningful research on clinical trials themselves, helping us understand the mechanisms of drug toxicity and trial failure – turning the existing, linear drug discovery paradigm into something much more effective – a system with feedback loops from human health to basic research.

Researchers will have access to a vastly larger and richer pool of data about patients, diseases, and treatment outcomes, as well as more sophisticated statistical services to analyze that data. Moreover, by sharing data, the research community can make more effective use of a limited



pool of patients. Every patient, every experiment contributes invaluable information (especially failures). It is essential that we capture this knowledge and widely disseminate it.

Ultimately, Health Commons enables the creation of new business models by replacing vertical integration with markets. New service providers will inevitably spring up to “lubricate” HC marketplaces for knowledge, services and resources. In addition to independent escrow agents, there will be directory services that help locate resources, match making services to locate partners, rating services to help comparatively evaluate quality and cost, services that simply integrate other offerings to provide a full-service solution, services that aggregate orders to reduce costs, service providers that partner with academic labs to make their state of the art tools broadly available as a service, FedEx-style logistic services that facilitate turnkey logistics and tracking so that outsourcing is as fast and safe (or even better) than doing everything within one institution, etc.

To realize the full potential, existing companies need to rethink their business models to leverage the commons. Here's an example. Only six out of the 1800 biotechnology companies funded since 1980 have made more money than was cumulatively invested in them. A key reason for the dismal record is that most biotechs, like most pharmas, are fixated on developing proprietary drugs. Many of these biotechs were originally funded because of their elegant platform technology e.g., a whole genome RNAi screening platform for validating genomic targets, or a proteomics platform capable of rapidly predicting how well potential leads bind with potential target proteins for a particular disease. Following the traditional business model, these companies used their platforms to identify a handful of targets and leads, then raised and invested tens or hundreds of millions in developing them, only to see them succumb to an “unpredictable” side-effect – one that might well have been predicted by another company's technology platform. How much more efficient it would be if biotech companies could be turned inside out, so that the real crown jewels -- their platforms -- were available as part of the Health Commons ecosystem, to anyone wishing to use them to find and validate targets and leads, in exchange for a share of the downstream revenue? What if investors could leverage an elegant platform technology into a piece of a hundred deals - undiluted by thousands of billable legal hours - instead of playing roulette with an unpredictable single-shot with one compound and low odds of success? Clearly, these advantages benefit all companies servicing the life sciences, from platform to discovery. The Commons, with its mix of technology and pre-negotiated agreements, makes this possible.

## **Health Commons transforms publishing into knowledge sharing.**

Scientific publishing is integral to the drug development process. But in the digital age, we must question whether the unit of a published paper is really the most efficient means of disseminating scientific knowledge. The elegance, clarity, and value of a carefully assembled, constructively peer-reviewed, professionally copy-edited and laid-out research article is clear. However, in this process, much information is delayed, or worse, lost. Interim data and results are typically discarded, especially the results of failed experiments, dooming others to waste time rediscovering them over again. Clinical trial data may never be published; a trial that fails because of an unknown toxicity, for which data has been captured previously, is both expensive --- and tragic for the patients involved. Although journals and funding agencies are committed, in principle, to requiring data associated with publications be made available, in practice, this only succeeds in the few cases for which community endorsed repositories exist. And beyond access to data, there's the deeper issue of making the conclusions conveyed in a scientific paper available in a structured form that can be understood and manipulated by computers as well as human scientists.

In Health Commons, all this will be different. By integrating the TOPAZ publishing platform, which currently supports PLoS ONE, into Health Commons, publication of research results will be a visible and automatically staged process. Knowledge will simply be promoted from one's personal repository in the Commons to be shared with one's laboratory, shared with one's collaborators, and ultimately to be made publically accessible. PLoS, along with other participating publishers, will provide vetting at many levels from community voting to review boards, as appropriate. Review by one's peers will occur at many stages: formal editorial boards could still provide traditional journal imprimaturs alongside more radical experiments in community voting. Peer review after publication will be implicitly based on how often a paper or piece of research is cited, and explicitly based on voting by applicable communities.

Supplementing traditional indexing of papers through PubMed, which can add delays as well as errors, papers will be semantically annotated by their authors using interactive Web Forms (faster, cheaper, and more accurate than PubMed's roomful of professional reviewers), then brought immediately to the attention of all researchers and research communities for whom that paper may be relevant based on inferences drawn using semantic models of the content of the paper, the domain of medical research, and the interests of the reader. And, of course, the full-text of these papers will be published under the Creative Commons attribution license, which means that anyone can develop services that enhance the value of the information.

Beyond ensuring timely access to knowledge by humans, semantic annotation is also the key to making that knowledge machine-understandable. The Health Commons platform will include basic applications that enable researchers not just to annotate papers, but comment on the annotations of others, and weave them into structured arguments that support and refute hypotheses. Capturing this structured discourse adds a new dimension to knowledge sharing not available in traditional publishing media. Moreover, other applications, developed by community members, can now use this structured knowledge, for example, to track and prioritize targets and leads, suggest experiments, or propagate knowledge selectively across communities. Because the range of potential applications is unlimited, computer access to published data and knowledge is likely one day to be at least as important as eyeball access. An early instance of such a future is Neurocommons, a major initiative to structure the world's knowledge of neuroscience for use by computers – which will become the first knowledge-sharing community within Health Commons.

### **From vision to reality**

Health Commons is a bold vision, whose aim is no less than to change the processes, economics and culture of translating basic research into drugs. Health Commons will transform academic biomedical research by making it cost effective for small groups of researchers to conduct industrial quality R&D. It will transform non-profit foundations into Virtual Biotechs, capable of developing therapies at a fraction of the time and cost of creating a traditional biotech, by leveraging an ecosystem of shared knowledge and services. Our hope is that pVirtual Biotechs can bring the cost of therapy development within the collective means of a few hundred patients with a rare disease, and produce a treatment within their lifetimes. Ultimately we expect Health Commons will transform pharma/biotech, as the internet has transformed nearly every other industry. It may even transform clinical medicine by making it possible to create Virtual Pharmas for individual patients – the ultimate in personalized medicine.

While the vision is bold, getting started is easy. Health Commons can be built incrementally, one service at a time, with the promise of significant impact in months rather than years. Every new collaboration built on the Commons will add new data and services to the ecosystem to the benefit of all. Many components already exist in the form of public data- and knowledge bases, computational models, contract research services, open source platform software and standardized legal agreements.

Of course there will be challenges convincing academic and industry competitors to cooperate for their own and the greater good. The very notion of a Health Commons challenges some core traditions of independent scientific research. First and foremost, it requires the commitment of previously independent scientists to work together to systematically elucidate the molecular biology of a disease, and to develop effective treatments. It demands a willingness to share information, results, specimens and resources, and occasionally to dynamically re-allocate resources; it requires a willingness to rethink traditional multi-year processes involved in moving research to the clinic by integrating and parallelizing the numerous steps involved in discovery and development. Last, it requires the continual coordination of efforts and results so that alternative paths can be prioritized and systematically explored.

The Health Commons is too complex for any one organization or company to create. It requires a coalition of partners across the spectrum. It is also too complex for public, private, or non-profit organizations alone - reinventing therapy development for the networked world requires, from the beginning, a commitment to public-private partnership. Only through a public-private partnership can the key infrastructure of the Commons be created: the investments in the public domain of information and materials will only be realized if that public domain is served by a private set of systems integrators and materials, tools and service providers motivated by profit. And in turn, the long-term success of the private sector depends on a growing, robust, and self-replenishing public domain of data, research tools, and open source software.

The direct beneficiary of the Health Commons is the consumer of health care - that is to say, all of us. We all have a significant personal incentive to accelerate the rate at which basic scientific research is translated into meaningful discovery and brought to the market, whether as new drugs or screens or simply a better understanding of the risks we face due to our genes and environments.

Please join our quest to make this vision a reality. Visit [HealthCommons.net](http://HealthCommons.net) for more information.