Available online at www.sciencedirect.com
SCIENCE DIRECT®
Research Policy 35 (2006) 839–853
research policy
www.elsevier.com/locate/respol

Inside the Anticommons: Academic scientists' struggle to build a commercially self-supporting human mutations database, 1999–2001

Stephen M. Maurer *

Goldman School of Public Policy, University of California at Berkeley, United States

Abstract

According to proponents of 'the tragedy of the Anticommons' hypothesis, resources with fragmented ownership are often underexploited. This paper describes how approximately 100 academic biologists tried to trade their data for corporate support of a worldwide Depository for human mutations data. Although the group obtained a $ 2.3 million offer, it became deadlocked in trying to respond. Analysis of this failure suggests that Anticommons effects exist in biotechnology. Deadlock occurred because most members could not afford the information costs needed to reach a decision. Project advocates tried to reduce the Anticommons effects by making information easier to obtain. Opponents added to the information burden of members by making assertions that could not easily be checked.

Keywords: Anticommons; Databases; Biotechnology; Research consortia; Technology transfer

1. Introduction

Despite recent increases in federal support for biomedical R&D, many worthwhile projects still go unfunded. This is particularly true for databases, which many government agencies tend to ignore in favour of more glamorous projects (Maurer et al., 2000). Given this shortage of tax dollars, scientists have become increasingly interested in schemes for tapping private sector resources. So far, however, private/public cooperation has been limited to charitable donations and quasi-charitable 'consortia'.[1]

Much more private money would be available if scientific communities were prepared to offer genuinely commercial, value-for-value exchanges. Biology databases are a case in point. Except for core data sets (e.g. gene maps), commercial firms usually find it prohibitively expensive to collect, rationalise, and unify academic data. This suggests that industry should be willing to pay academic biologists to create better data sets in the first place. This paper describes how the Mutations Database Initiative (MDI) – a worldwide organisation of academic scientists who study human mutations – tried to cre-

* Correspondence to: 2632 Hilgard Ave., Berkeley, CA 94709, United States. Tel.: +1 510 848 3593; fax: +1 510 643 9657.
E-mail address: maurer@econ.berkeley.edu.

[1] During the 1990s, commercial biotechnology and pharmaceutical firms repeatedly engaged in well-publicised high-profile 'partnerships' and 'consortiums' with academic scientists. Despite this terminology, none of these efforts constituted a genuinely commercial exchange of value. Instead, corporations paid academic biologists to do work that they wanted to do anyway. Probably the best-known example involved the so-called 'SNP Consortium,' in which large pharmaceutical companies paid academic scientists to place data in the public domain so that it could not be patented. The companies believed that they would pay dearly if SNPs were patented.

0048-7333/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.respol.2006.04.008

ate an advanced technology, self-supporting database by negotiating an agreement with commercial partners. The account is based on the author's experiences as a practising lawyer who participated in negotiations between 1999 and 2001.[2]

The idea of making large academic biology databases commercially self-supporting is so promising that it is natural to ask why it has not happened already. The principal obstacle seems to be that any proposed institution would have to receive active support from a large fraction of the community's roughly 100 database providers. The hypothesis known as 'The Tragedy of the Anticommons'[3] states that multi-party consent is extremely difficult to obtain and that resources that can be vetoed by small minorities are systematically underused. Anticommons effects are predicted to be significant in biotechnology and pharmaceutical research (Heller and Eisenberg, 1998), although survey evidence has thus far failed to detect them (Walsh et al., 2003). This paper provides evidence that Anticommons effects prevented human mutations biologists from exploiting commercial resources to build a new facility that would have unified and rationalised their work.

Section 3 describes the mutations community and the database problems it set out to solve. Section 4 describes the institution that the community planned to build. Section 5 describes how the community tried and ultimately failed to create this institution. It also recounts how individual actors struggled for control of the facility. Section 6 uses older industrial organisation theories based on transactions costs, bounded rationality, and information impactedness to interpret the mutations community's experience. The handful of well-informed actors struggling for control of the facility tried to increase or alleviate these conditions depending on their interests. Section 7 asks what lessons scientists and governments should draw from the mutations community's successes and failures. Section 8 presents a brief conclusion.

2. Background: the community and its problem

2.1. The problem

Biologists spent most of the 1980s and 1990s recording the DNA 'letters' locked inside the human genome. Today, the challenge is to understand what the letters mean; i.e. to discover how each portion of the genome controls processes within our bodies. Mutations data are one important clue. Hospital patients who suffer from certain diseases are routinely tested for genetic abnormalities. Comparing abnormalities against patient symptoms helps biologists decode the messages written in the genome.

As of 2000, most of the world's mutations data was found in roughly 100 Web-based locus specific databases (LSDBs), each of which specialised in between 1 and 10 locations in the human genome. Because of this narrow specialisation, LSDB data was carefully edited and had very high quality. Furthermore, at least half of the data recorded in LSDBs had never been published anywhere else. This fraction was sure to rise as biology journals became more and more selective about publishing data (Cotton, 2000).

However, LSDBs also had important weaknesses. First, most LSDBs were small, fragmented, and employed non-standard computing conventions. This made it difficult for scientists to combine and compare data from different genes. The situation was aggravated by the fact that individual LSDBs sometimes applied different nomenclatures, reported different categories of information, and used different quality control ('curation') practices. Second, most LSDBs relied on primitive computing architectures and search tools. This made them much less powerful than contemporary commercial databases. Third, LSDBs were incomplete. Informal surveys showed that most hospital clinics and private laboratories did not report their discoveries to any database. Finally, there were not enough LSDBs to go around. Existing LSDBs covered approximately 270 genes. This was less than 1% of the roughly 35,000 disease- and cancer-causing mutations believed to exist (Horaitis, 2000a).

A few groups had developed central databases to meet these problems. This strategy had only limited success. First, no one had been able to handle the flood of data contained in LSDBs. Instead, all existing databases had adopted a 'mile wide, inch deep' strategy that barely scratched the surface of what LSDBs had to offer. Sec-

[2] Clearly, it would have been better for an outside observer to tell the story. Since no such person exists, I have tried to do the next best thing by limiting my narrative to specific facts documented in my collection of more than 3000 pages of e-mails, memos, letters and other materials related to the project. Re-reading these documents has helped me to refine – and in some cases to revise – my contemporaneous understanding of why things happened or what particular actors wanted to accomplish. In a few cases I have relied on memory for essential conversations or meetings. On these occasions I have limited myself to objective information concerning who said what, and when.

[3] The term 'Anticommons' is a play on words and refers to the 'Tragedy of the Commons' which is usually taught in freshman economics. For the locus classicus of this popular illustration of what economists recognise as a 'common pool problem', see Hardin (1968). The idea of an 'Anticommons' is based on the notion that property rights can also *destroy* assets by promoting friction and deadlock (Heller and Eisenberg, 1998).