

SCRIPT-ed

Volume 1, Issue 4, December 2004

In Code, We Trust? Regulation and Emancipation in Cyberspace

*By Zhu Chenwei**

Abstract

Code is one of the regulatory modalities as identified by Lawrence Lessig. It is proposed that, in cyberspace, code should not only regulate but also emancipate. However, the emancipatory dimension of code has long been neglected since law and market are increasingly operating in a normative vacuum. The emancipatory approach is also supported by the practice of digital commons, which is to liberate cyberspace from various constraints.

DOI: 10.2966/scrip.010404.585

© Zhu Chenwei 2004. This work is licensed through [SCRIPT-ed Open Licence \(SOL\)](#).

* ZHU Chenwei, LL.B. (SISU, Shanghai), LLM in *Innovation, Technology and the Law* (Edinburgh). He is currently a member on the team of iCommons China. The Author is thankful to Mr. Andres Guadamuz, who always provides rich inspiration and great food for thought. This article is impossible without some fruitful discussions with Gerard Porter, David Possee, Isaac Mao, Sun Tian, Chen Wan, Michael Chen and Zheng Yang. The author is also indebted to the anonymous reviewers providing invaluable comments on this paper. He is solely responsible for any mistake in it.

Ronald Dworkin:

We live in and by the law. It makes us what we are: citizens and employees and doctors and spouses and people who own things. It is sword, shield, and menace [...] We are subjects of law's empire, liegemen to its methods and ideals, bound in spirit while we debate what we must therefore do. (*Law's Empire*¹, 1986)

Lawrence Lessig:

We live life in real space, subject to the effects of code. We live ordinary lives, subject to the effects of code. We live social and political lives, subject to the effects of code. Code regulates all these aspects of our lives, more pervasively over time than any other regulator in our life. (*Code and Other Laws of Cyberspace*², 1999)

1. Introduction: Deciphering the Lessigian Code

On 3 February 1776, Samuel Johnson wrote to James Boswell, an Edinburgh lawyer as well as his most loyal and learned friend, claiming that “[l]aws are formed by the manners and exigencies of particular times, and it is but *accidental* that they last longer than their causes.”³

If Dr. Johnson was right, today's rapidly changing information and communication technologies (ICTs) could subject the law to an even more volatile environment. Law is undergoing an ever-growing legitimization crisis: it was originally designed to cushion blows of social changes and thus bring certainty for social progress. The “accidentalness” of law, however, hardly lets this sweet expectation be fulfilled.

As for regulating cyberspace, Lawrence Lessig proposed that the “accidentalness” of law could be remedied by adopting “code,” because “code is law” when put in the new techno-social context. Why is code of such significance for cyberspace? Prophetic Cyberpunk novelists, such as William Gibson, articulated a vision that cyberspace represents “an effective and original symbol of human involvement of machines.”⁴ The rhythms of everyday life are increasingly dominated by the clanking world of machines in which technical code rules. In this scenario, Lessig urges us to “look to the code” by stressing that “we must build into the architecture a capacity to enable choice—not choice by humans but by machine.”⁵

Following Lessig's logic, two crucial questions now spring before us: (1) Is it always possible to design code with architecture that has the capacity to “enable choice”? (2)

¹ Ronald Dworkin, *Law's Empire* (London: Fontana, 1986) p. vii.

² Lawrence Lessig, *Code, and Other Laws of Cyberspace* (New York: Basic Books, 1999) p.233.

³ James Boswell, *The Life of Samuel Johnson, LL.D.* (Hertfordshire: Wordsworth Editions Ltd., 1999) p. 472; (emphasis added).

⁴ See Paul Alkon, “Deus Ex Machina in William Gibson's Cyberpunk Trilogy”, in George Slusser and Tom Shippey (eds.) *Fiction 2000 : Cyberpunk and the Future of Narrative* (Athens; London: University of Georgia Press, 1992) p.76.

⁵ Lessig, *Code and Other Laws of Cyberspace* (New York: Basic Books, 1999) p.163.

If the answer is yes, *how* can we design this ideal-typed architectural framework for cyberspace?

David Post was dubious about Lessig's approach. He agreed that the value-laden architecture was important for decision-making, but he did not accept that this structure called for "more politics to help devise the plan for the codes of cyberspace."⁶ He believed that more emphasis should be placed on liberty rather than on collective power and thus accused Lessig of creating a big "Liberty Gap"⁷ between cyber-libertarians and technological determinists.

Interestingly, Post's critique of the Lessigian code might fall into the agenda of the so-called New Chicago School, named by Lessig himself. In 1998, Lessig, for the first time, identified the four regulatory modalities—law, market, architecture, and norm—for legal analysis under the name of the "New Chicago School,"⁸ which was recognized in a somewhat informal way by his colleagues at the University of Chicago.⁹ This School has no unified methodologies, nor any unified normative implications. What unifies the New Chicago School is the subject matter: "the ways that law can influence behavior *indirectly*, by changing social norms."¹⁰ The approach is at least logically consistent with David Post, who urges us to focus less "on trying to specify optimal configurations of legal systems," and more on "the best *algorithm* for finding more acceptable rule configurations."¹¹

It is also noteworthy that there is a grave danger when embracing the New Chicago School, because the School has a totalizing agenda to colonize every aspect of life under regulation—or more precisely, under power and control. Lessig said this was obviously the "dark side" of the approach.¹² This article is about the regulation of cyberspace, but it is written in an attempt to avoid the aforementioned "dark side".

Code is both *regulatory* and *emancipatory* for cyberspace. As legal theorist Boaventura de Sousa Santos points out, modernity is grounded on "a dynamic tension between the pillar of regulation and the pillar of emancipation."¹³ The discovery of this long-ignored *emancipatory* dimension of code throws new light on the study of cyberspace regulation.

⁶ David G. Post, "What Larry Doesn't Get: Code, Law and Liberty in Cyberspace", (2000) 52 *Stanford Law Review* 1439 at 1458.

⁷ *Ibid.*

⁸ See Lessig, "The New Chicago School", (1998) 27 *The Journal of Legal Studies* 661.

⁹ The University of Chicago Law School Roundtable arranged an interdisciplinary program series for discussing the New Chicago School. See, especially, "The New Chicago School: Myth or Reality", (1998) 5 *The University of Chicago Law School Roundtable* 1.

¹⁰ According to Eric Posner, the self-proclaimed School uses a number of intellectual tools. For example, Posner himself uses game theory, Randal C. Picker uses computer-generated models and Tracey L. Meares uses sociology etc. see *Ibid.* (emphasis added).

¹¹ See David G. Post and David R. Johnson, "'Chaos Prevailing on Every Continent': Towards a New Theory of Decentralized Decision-making in Complex Systems" (1998) 73 *Chicago-Kent Law Review* 1055, at 1092 (emphasis added).

¹² Lessig, *supra* note 9, at 691.

¹³ Boaventura de Sousa Santos, *Toward A New Legal Common Sense: Law, Globalization and Emancipation* 2nd ed. (London: Butterworths LexisNexis, 2002) p.2.

Figure 1: A Cognitive Map of Regulation/Emancipation in Cyberspace

REGULATORY MODELS	THE ROLE OF CODE IN REGULATING CYBERSPACE	THREE CHOICES OF INFORMATION POLICY (YOCHAI BENKLER)	METAPHORS /ANALOGIES	EXAMPLES
Code	Code	N/A (The Net Has Not Been Commercialized)	Garden of Eden	ARPANET
Code + Market CM Model	Code is Led by Market	Self-Regulation	Invisible Hand	Search Engine Market
Code + Law CL Model	Code is Led by Law	Government Intervention	Second Enclosure Movement (<i>James Boyle</i>)	Content Industry (DRM, Copyright and Copyright Levies)
(Code + Market) + Norm (Code + Law) + Norm Commons Model	Code's Normative Value Released	Intellectual Commons	Information Eco-System	Wiki Projects (<i>Wikipedia</i>)

Sources include Yochai Benkler,¹⁴ and James Boyle.¹⁵ The Metaphors/Analogies section is inspired by Prof. Alfred C. Yen's incisive analysis of two metaphors of cyberspace.¹⁶

Code liberates and constrains, but it is not, nor should it be, an excessively dominant element. It is fruitful only when it enjoys full interplay with the other three regulatory/emancipatory modalities (See *Figure 1*). This article analyzes three kind of regulatory mechanisms respectively in their dominating areas. **(1) Co-regulation of Code and Market (CM Model):** This model prevails in the burgeoning search engine market. A focus will be put on the leader of the industry: Google Inc. **(2) Co-regulation of Code and Law (CL Model):** In this model, the content industry is struggling to weave a seamless web of controlling their copyright works by using both code (like Digital Rights Management) and law (such as anti-circumvention law or imposition of copyright levies). A gloomy prospect lies ahead when the benefits brought by technological progress could be effectively effaced by a strict control of Law and Code. **(3) Introduction of Norms to the above two models (Commons Model):** Strictly speaking, a norm does not *regulate*: but *liberate*. Norms insert an emancipatory element into the mechanism of "regulation." This mechanism is exemplified by wiki projects, which enable collaborative web editing on a grassroots level.

¹⁴ See, generally, Yochai Benkler, "The Commons as a Neglected Factor of Information Policy", 26th annual Telecommunications Policy Research Conference, Oct 3-5, 1998 available @ <http://www.benkler.org/commons.pdf>

¹⁵ James Boyle, "The Second Enclosure Movement and the Construction of the Public Domain", (2003) 66 *Law and Contemporary Problems* 33.

¹⁶ Alfred C. Yen, "Western Frontier of Feudal Society?: Metaphors and Perceptions of Cyberspace" (2002) 17 *Berkeley Technology Law Journal* 1207.

Reidenberg stresses that the Net should be recognized as a collection of “semi-sovereign entities”¹⁷ that is self-sustainable and self-supportable. However, the Internet is not immune to the danger of being an “Electronic Babel Tower,”¹⁸ with unending quarrels and profound misunderstandings among profit-hungry and power-thirsty parties struggling for more and more control. The emancipatory model—the Commons Model—also relies upon Code and the other three elements, and would force these parties to re-liberate the Net from a heavily regulated and controlled environment.

2. In Search of Justice: A Taste of Googлизм

The search engine market is a brave new world. It remains one of the few realms that has not yet been fully colonized by law from the off-line world. On another front, norms are also slow to take shape to address the ethical concerns in this area. For these reasons, it is an ideal model upon which to consider the effect of the first regulatory mechanism—the Code + Market (CM) Model.

2.1 When Code is Marketable...

Search engines embody the dream of “pull” technology, in contrast to a “push” media delivery model. The push model entails “the process of automating the searching and delivery of content to a user’s PC”¹⁹ without a request being made. The pull model lets users decide what they may need following a “pull”²⁰ and then issues tailored relevant information.

On the immense digital seas, the pull model, represented by search engine technology, has proved to be a desirable way for delivering digital information, and has already achieved significant commercial success. The rising star in the kingdom of search technology is an oddball company—Google.com—that seems quite good at surprising, if not irritating, the old business world.

In 1998, two Stanford PhD students, Sergey Brin and Lawrence Page, started their business (originally a research project) in a university dorm room.²¹ The name Google was playfully derived from the word “googol.” In 1938, “googol” was coined by Milton Sirota, when his uncle, a famous mathematician named Edward Kasner, asked him “what name he would have given to a ‘1’ followed by 100 zeros.” The term “googol” then came into existence as the number 10 raised to the power 100 (i.e. 1

¹⁷ Joel R. Reidenberg, “Governing Networks and Rule-Making in Cyberspace”, (1996) 45 *Emory Law Journal* 912 at 930.

¹⁸ Egbert J. Dommering, “Copyright Being Washed away through the Electronic Sieve: Some Thoughts on the Impending Copyright Crisis”, in PB Hugenholtz (ed.) *The Future of Copyright in a Digital Environment* (The Hague; London: Kluwer Law International, 1996) p.11.

¹⁹ See Paul Chin, “Push Technology: Still Relevant After All These Years?”, *Intranet Journal—The Online Resource for Intranet Professionals*, 23 July 2003, available @ http://www.intranetjournal.com/articles/200307/pij_07_23_03a.html

²⁰ See the definition given by *Webopedia*, “Pull”, @ <http://www.webopedia.com/TERM/P/pull.html>

²¹ Peter Jennings, “The Google Guys—Larry Page and Sergey Brin Founded World’s Most Popular Search Engine”, *ABCNEWS.com*, 20 February 2004, available @ http://abcnews.go.com/sections/WNT/PersonofWeek/pow_google_guys_040220.html

googol = 1.0×10^{100}).²² “Googol” thus betrays the ambition of Brin and Page, who venture to tap the infinite potential of the Internet through employment of their powerful algorithms.

On 19 August 2004, Google Inc. went public became a listed company on the Nasdaq (marked as GOOG) with a debut price of \$85 a share.²³ The success owes much to an innovative algorithm called PageRank, which was granted a patent (US Patent 6,285,999) in 2001.²⁴ Google uses PageRank to rank websites in search results.²⁵ Early search engines did not provide an ordering of results. The code owned by Google is brilliantly written, but it is far from perfect—it is not difficult for sophisticated webmasters to abuse the code. In a brutal marketplace, a large number of companies, small and big, are highly motivated to manipulate search engines like Google by employing imaginative “cheating” techniques. It is difficult to predict the direction in which the market forces are leading Google’s code.

2.2 Dancing with Spamdexers

The search engine business is the crown jewel of today’s eyeball economy. Unlike many other self-dramatizing dot-com companies, Google.com does actually make a lot of money. In 2003, Google received a net income of \$105.6 million,²⁶ with ninety-six percent of this revenue coming from advertising.²⁷ In an attempt to capitalize on searching technologies, Google launched its lucrative paid search service, known as “sponsored links,” which appear on the far right-hand side of search results. Google continually develops promotion programs for advertisers to bid for their ranking in sponsored links,²⁸ which means the “relevance” is *partly* decided by how much was paid to Google. Google, though, was not the first to sell search service; Yahoo’s Overture did this earlier.²⁹

²² See “Milton Sirotta” in *Encyclopedia of WordIQ.com*, @ http://www.wordiq.com/definition/Milton_Sirota.

²³ See “Google Set for Market Debut”, *Reuters UK KNOW. NOW*. 19 August 2004 @ <http://www.reuters.co.uk/newsPackageArticle.jhtml?type=businessNews&storyID=567943§ion=news>.

²⁴ See the website of United States Patent and Trademark Office, *USPTO Patent Full-Text and Image Database*, September 4, 2001 <http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=/metahtml/search-bool.html&r=1&f=G&l=50&col=AND&d=ptxt&s1=6,285,999.WKU.&OS=PN/6,285,999&RS=PN/6,285,999>.

²⁵ Phil Craven, “Google’s PageRank Explained and How to Make the Most of It”, *WebWorkShop* @ <http://www.webworkshop.net/pagerank.html>.

²⁶ *RedHerring*, “IPO Watch: Dot-Com Math”, 16 August 2004 @ <http://redherring.com/Article.aspx?a=10795§or=Capital&subsector=PublicMarkets&hed=IPO+Watch%3A+Dot-com+math>.

²⁷ Andrew Orlowski, “Google Files Coca Cola Jingle with SEC”, *The Register*, 29th April 2004, http://www.theregister.co.uk/2004/04/29/google_s1_filed/.

²⁸ *SiliconVally.com*, “Google Will Let Web Sites Bid for Higher Rankings In Ad Listings”, 19 February 2002 @ <http://www.siliconvalley.com/mld/siliconvalley/news/editorial/2704974.htm>.

Google will let Web sites bid for higher rankings in ad listings.

²⁹ See *The Economist*, “How Good is Google?” Oct 30th 2003, available @ http://www.economist.com/displaystory.cfm?story_id=2173573.

It is worried that the pay-for-placement model put the search engine sector gradually on a morally slippery downward slope. For example, Yahoo!'s paid search service is even more crafty than Google's. Instead of selling keyword directly, one of Yahoo!'s advertising strategies involves selling the "updatedness" of search results: traders pay to have their website re-crawled by spiders (or robots, web crawlers) every 48 hours to be listed in Yahoo's search database³⁰ (not their directory, which is manually made³¹). Otherwise, websites have to wait for weeks before Yahoo!'s spiders re-visit. If we regard "updatedness" as a *temporal* dimension of "relevance," this paid inclusion policy indicates that the search results are determined partly by one's financial power. Those who pay regularly will, of course, have a more up-to-date presence in search results.

The online auction giant eBay is another striking example as to how search results might be "distorted" by the gravity of huge market power. In the summer of 2003, eBay sent a letter to Google, requesting Google to prohibit advertisers from using eBay's trademark in the "sponsored links" service. They also attached a 13-page long list detailing a series of eBay-forbidden words (e.g. "eBay Power Seller"). eBay's requested ban might stifle competition by arbitrarily excluding all eBay management software from the paid search service.³²

A few months later, eBay was found to be involved in spamming search engines by using "cloaking" techniques (the term will be explained in detail in the next paragraph). The eBay Affiliate Program encourages individuals to use various means to deliver traffic to eBay.³³ Ryle Goodrich, an over-enthusiastic participant to the Program, used his own website as a doorway to "steer Google search[s] over to eBay."³⁴ He wrote a computer program to "harvest auction description[s] from eBay to fill his Web pages with terms and content to attract Google's spiders."³⁵ Around 500,000 pages were affected and redirected to eBay sites by Goodrich's "innovative" program.

Cloaking, however, is not a new thing. Many search engine optimization (SEO) companies have used this trick for years.³⁶ There is even commercial cloaking software on the market, at prices from USD \$25 to over \$1,000.³⁷ Not all cloaking is unethical. Initially cloaking was used to feed search engines, which lacked the

³⁰ *The Economist*, "Spiders in the Web", *A Survey of E-Commerce*, May 15th 2004, p.16.

³¹ A Yahoo search database uses spiders, but its directory is still manually made. The difference is highlighted by Shawn Campbell, "What is Site Match?" *SiteProNews* May 17, 2004 @ <http://www.sitepronews.com/archives/2004/may/17prt.html>.

³² See Stefanie Olsen, "Google Ads a Threat to eBay Trademark", 8 August 2003, *CNET News* @ http://news.com.com/2100-1024_3-5061888.html?tag=fd_top.

³³ See the webpage of eBay Affiliate Program @ <http://affiliates.ebay.com/>.

³⁴ Ina & David Steiner, "Google Shuts Off Participant in eBay Affiliate Pilot Program", *AuctionBytes* 24 October 2003 @ <http://www.auctionbytes.com/cab/abn/y03/m10/i24/s00>.

³⁵ Ina & David Steiner, "Battle for Eyeballs Drives Google Traffic to eBay", *AuctionBytes* 24 October 2003 @ <http://www.auctionbytes.com/cab/abn/y03/m10/i24/s01>.

³⁶ See Danny Sullivan, "Bait-And-Switch Gets Attention", *SearchEngineWatch* 2 June 1999 @ <http://searchenginewatch.com/sereport/article.php/2167201>.

³⁷ See Michelle Anderson, "Cloaking and Page-Jacking" @ <http://www.mailbot.com/articles/titles/2.html>.

capability to crawl and index a dynamically updated page by providing IP delivery of a “static” page as an alternative.³⁸ Later on, webmasters found that cloaking could be used to boost their site ranking. A “cloaker” delivers one version of a page to ordinary users, and a different version (possibly a stolen page from its competitors) to spoon-feed search engines.³⁹ For example, State Farm Insurance once submitted the domain name www.Statefarm.com especially for search engines, while ordinary people only see their homepage at www.Statefarm.com.⁴⁰ In another infamous case, the SEO company Green Flash helped to “pagejack” its client’s website (Data Recovery Labs) by stealing its competitor’s pages. Though the case was filed before the U.S. Federal Trade Commission (FTC), the authority failed to take an immediate action.⁴¹

Cheating search engines through cloaking is only the tip of the iceberg of spamdexing (search engine spamming), which refers to “the practice on the World Wide Web of deliberately modifying HTML pages to increase the chance of being placed high on search engine relevancy lists.”⁴² Spammy websites are equipped with a surprising number of ways to cheat. They can “spamvertize” (advertise by spamming) themselves by whatever means possible. In order to fight increasingly rampant spamdexing, Google specifies seven commandments against possible abusive practices in its *Webmaster Guidelines*:

- Avoid hidden text or hidden links.
- Don't employ cloaking or sneaky redirects.
- Don't send automated queries to Google.
- Don't load pages with irrelevant words.
- Don't create multiple pages, subdomains, or domains with substantially duplicate content.
- Avoid "doorway" pages created just for search engines, or other "cookie cutter" approaches such as affiliate programs with little or no original content.⁴³

2.3 PageRank and Its Algorithmic Weakness

Unfortunately, Google’s above list is far from comprehensive. So far the Google team has failed to cover at least two of the most “cutting-edge” spamdexing techniques—wikispamming and blogspamming.

Wikispammers specialize in boosting their ranking in search results by inserting links of their spamvertized pages in large amounts of wiki pages. A wiki is “a piece of

³⁸ See Greg Boser, “Green Flash Pagejacking Evidence”, @ <http://www.webguerrilla.com/pagejacking/>.

³⁹ See *SearchEngineWorld*, “Cloaking Overview”, @ <http://www.searchengineworld.com/index.htm>.

⁴⁰ Danny Sullivan, “A Bridge Page Too Far?” *SearchEngineWatch* 3 February 1998 @ <http://www.searchenginewatch.com/sereport/article.php/2165891>.

⁴¹ Danny Sullivan, “Pagejacking Complaint Involves High-Profile Sites”, *SearchEngineWatch.com*, 12 May 2000 @ <http://searchenginewatch.com/sereport/article.php/2162601>.

⁴² *Wikipedia*, “Spamdexing”, @ <http://en.wikipedia.org/wiki/Spamdexing>.

⁴³ Google, *Webmaster Guidelines*, @ <http://www.google.com/webmasters/guidelines.html>.

server software that allows users to freely create and edit Web page content using any Web browser.”⁴⁴ Blogspammers do virtually the same thing by “repeatedly placing comments to various blog posts that provided nothing more than a link to the spammer’s commercial web site.”⁴⁵

Here we are posed an interesting question: what has inspired these spammers to do the wicked things? The author believes that the following two elements make spamming possible: (1) the market is an irresistible force that motivates advertisers to develop cheap way of promoting themselves such as bombarding wikis and blogs. (2) Spamdexers want to take advantage of the algorithmic weakness of Google’s code—PageRank—which is responsible for ordering the search results.

A further examination into PageRank’s working mechanism helps to clarify the issue. Generally speaking, PageRank is a ballot system, which ranks the popularity of a website by calculating the number of its links appearing on other websites. It is said that Google uses PageRank to mirror public opinions.⁴⁶ For example, if website A has a link to website B, this means that A electronically gives B a vote, because the A believes that B is good enough to be link-worthy.⁴⁷ The wider a website is externally linked by other sites, the more visible (i.e. relevant) this site will be, the higher it will be placed in search results. In other words, Google interprets relevance in terms of link popularity.

Let’s take a more concrete example. For this I used *Link Popularity Checker*, a program provided by SeoToolkit.co.uk, to calculate the number of external links for three UK full-text online legal journals: the *Journal of Information, Law and Technology (JILT)*, the *Web Journal of Current Legal Issues (WJCLI)*, and *SCRIPT-ed*. As mentioned above, the relevance of a Google search will be sorted by link popularity ranking. *Figure 2* shows how widely each journal is linked, i.e. how many websites vote for the journal with their links. It is not difficult to see that *JILT* is currently the most widely linked, and thus the most popular of the three. If a reader does a Google search (possibly with a combination of keywords of “Online + UK + Law + Journal”), she will find that *JILT* has a higher position over the rest two,⁴⁸ because it enjoys the highest link popularity and is seen by Google as the most “relevant.”

Figure 2: “Vote Me by Providing a Link to Me, Please!”-- Link Popularity Checker

URL OF THE ONLINE JOURNALS	ALL THE WEB	ALTA VISTA UK	GOOGLE/AOL UK	MSN UK	YAHOO	TOTAL:
----------------------------	-------------	---------------	---------------	--------	-------	--------

⁴⁴ For more information about wiki technology, see Wiki.org, “What Is Wiki” @ <http://wiki.org/wiki.cgi?WhatIsWiki>.

⁴⁵ Wikipedia, “Spamming”, @ <http://en.wikipedia.org/wiki/Spamming>.

⁴⁶ Chris Beasley, “What is Google—Watch?”, @ <http://www.google-watch-watch.org/>.

⁴⁷ See Google.com, “Google Searches More Sites More Quickly, Delivering the Most Relevant Results”, @ <http://www.google.com/technology/>.

⁴⁸ The search can be done manually or by using software such as “Free Monitor for Google”, downloadable @ <http://www.cleverstat.com/google-monitor-query.htm> free of charge. This Monitor can display many keywords and resulted positions of certain URLs ranked Google searches in a very effective way.

http://elj.warwick.ac.uk/jilt/ (<i>Journal of Information, Law and Technology</i>)	1860	1900	332	0	2180	6272
http://webjcli.ncl.ac.uk/ (<i>Web Journal of Current Legal Issues</i>)	842	852	180	1376	1000	4250
http://www.law.ed.ac.uk/ahrb/script-ed/ (<i>SCRIPT-ed</i>)	287	294	50	0	306	937

Source: The results were automatically generated by running Link Popularity Checker, available at <http://www.seotoolkit.co.uk/link_popularity_checker.asp> on 21 August 2004. This research can also be done manually. If you use Google, enter the command “*link:http://example.com*” in the search box, the number of links directed to it will be displayed.

It is worth mentioning at this point that PageRank does not count all links. Google struggles to exclude links placed solely for maximizing rankings. For example, an early cheating technique involved using “link farms,” which are sites solely dedicated to stuffing hyperlinks into their websites. Link farms improve search ranking for other sites by artificially increasing the number of links.⁴⁹ Google sees link farming as disturbing and distorting, and is unforgiving in punishing sites playing this trick. Sites affiliated with link farms risk being removed from Google’s index.⁵⁰

In contrast to link-farming today, wikispammers and blogspammers play in another large grey area of which the Google team is not yet fully aware. These “innovative” spammers incorporate publicly editable wikis and blogs with open comment systems into their new spamdexing bases. Spamming can be either done manually or by deploying spambots that crawl wikis and blogs and leave links automatically.

In May 2004, spambots attacked Meatball Wiki, a prominent inter-community dealing with online culture.⁵¹ Wikizens have taken great efforts to fight spammers since then. The task of blocking spammers’ IP address, along with fighting other forms of “wiki vandalism,” is becoming a daily and sometimes hourly routine of wiki administrators. Most wiki page writers delete the spammed pages and reverse them to earlier “clean” versions. However, deletion does not necessarily hurt spammers’ ranking in search results, because wiki projects generally archive earlier changed history page—including the spammed ones. Google’s spider would still find and index these “dirty” pages.

Besides wikizens, bloggers were also under massive attack at roughly the same period of time. Almost all bloggers using Movable Type (MT), a famous personal e-publishing system, suffer from repeated nasty advertising posts in their comments. Spammers detected and found loopholes in the code of MT software and manage to

⁴⁹ For a detailed discussion of link farms, see Kim Krause, “Warning!! Link Farm Ahead”, *SearchEngineGuide.com*, December 19, 2002 @ http://www.searchengineguide.com/krause/2002/1219_kk1.html.

⁵⁰ See Larisa Thomason, “Link Farms Grow Spam”, *Webmaster Tips Newsletters*, April 2002 @ http://www.netmechanic.com/news/vol5/promo_no7.htm#.

⁵¹ *MeatBall News*, “MeatballWiki and UseModWiki get WikiSpammed by Bots”, 22 May2004, @ <http://www.usemod.com/cgi-bin/mb.pl?MeatballNews>.

turn personal weblogs into their link farms. Spamming was only partly controlled after MT-Blacklists were released for MT blog users.⁵²

Who then should be responsible for, and bear the cost of, spamdexing? At first sight, spammers are condemnable as a major source of evil. However, it is also believed that the appearance of spamdexed sites reflects the *algorithmic weakness* of a search engine.⁵³ Google tries to change the algorithm of the PageRank to fight against spammers, but is unable to constantly refine it.⁵⁴ In 2002, Search King, an SEO vendor, challenged Google's PageRank in an Oklahoma district court, alleging that Google's efforts to prevent spamdexing were arbitrary and unfair. Search King thought that its SEO service of increasing clients' visibility was perfectly legitimate and compatible with the PageRank standard. In 2003, the court ruled in favor of Google.⁵⁵ However, a critical question remains unanswered: does the ruling indicate that PageRank was a vulnerable technical standard that could not be properly implemented without human supervision? SEO practices distort page rankings, but should the Google team be empowered to remove whatever they believe to be inappropriate from the search list?⁵⁶ In fact, Google only has an obscure exclusion policy that allows it to remove pages on a seemingly *ad hoc* basis:

*Your page was manually removed from our index, because it did not conform with the quality standards necessary to assign accurate PageRank. We will not comment on the individual reasons a page was removed and we do not offer an exhaustive list of practices that can cause removal.*⁵⁷

What does Google mean by its "quality standards necessary to assign accurate PageRank"? It seems that the Google team wishes to be tight-lipped on this standard and lacks interest in making it transparent.

2.4 The Trilemma of the CM Model

Google is growing big and strong, and its power often goes unchecked. In a marketplace where the law of the jungle prevails, Code is made to be a humble servant to Market. The search engine world exemplifies a trilemma faced by the CM regulation model, with which both Code and Market operate in a normative vacuum.

⁵² For more information about the issue in progress, please see discussions in the MT-Blacklist Forum: @ http://www.jayallen.org/comment_spam/forums/index.php?s=9cac46b78bb7400186c182a802ccf8d3&act=idx.

⁵³ F. Gregory Lastowka, "Search Engines, HTML, and Trademarks: What's the Meta For?" (2000) 86 *Virginia Law Review* 835 at 865.

⁵⁴ See Andrew Orlowski, "Google's Ethics Committee Revealed", *The Register*, 17th May 2004, @ http://www.theregister.co.uk/2004/05/17/google_ethics_committee/.

⁵⁵ *Search King, Inc. v. Google Technology, Inc* Case No. CIV-02-1457-M.

See Notice of Order or Judgments Fed R. Civ. P. 77(d) available @ <http://research.yale.edu/lawmeme/files/SearchKing-SummaryJudgement.pdf>.

⁵⁶ Rob Sullivan, "Google sued over PageRank", January 10, 2003 *SearchEnginePosition.com* <http://www.searchengineposition.com/info/Articles/GooglevsSearchKing.asp>.

⁵⁷ Google, "My Web Pages Are Not Currently Listed", @ <http://www.google.com/webmasters/2.html>.

Firstly, Code in search engine world plays a double role: it enables people to have a much wider reach to Internet content with an ordered arrangement, while at the same time it constrains people from knowing what is outside search lists. Code such as PageRank is revolutionary and helps to open wider horizons, but that same Code can also be manipulated to produce a distorted worldview biased towards targeted commodities promoted by SEO vendors. To put it more directly, one code might be pitted against another code, and the turf-war between them could be endless and unfruitful. For example, some search engines want to control spamdexing by refusing doorway pages using fast meta-refresh techniques,⁵⁸ but SEO companies are “clever” enough to bypass this by employing new cheating methods, such as “code swapping” tricks.⁵⁹

Secondly, the development of Code is increasingly profit-oriented under market pressure and the market may lead code astray. Critics assert that a free market system has no way “for identifying common human needs,”⁶⁰ and only fills people with boundless vanity and drives them greedy and mad. Google was once an unconventional company with an informal and playful working culture. *The Register* reported that Google did not publish any of its research programs, but that it did publish *Google Daily Menus* dedicated to what is to be eaten by Google guys.⁶¹ Now that Google has gone public, we see that the Market is obviously not a democratizing force. The power to influence a company’s decision-making process is still concentrated in the hands of the few. The voting structure is as follows: stocks held by Google’s employees carries ten votes, while publicly traded stock represent only one vote, i.e. the executive management team and directors control 61.4 per cent of the voting power.⁶²

Google’s way of doing business is undeniably innovative, but this does not mean that the Google team can be exempt from ethical concerns. In April 2004, Google advanced a new project to the world: they provided Google Mail (Gmail) account whose storage is as big as one Gigabyte. The really new thing is not its huge storage capacity, but that Gmail uses robots to crawl one’s email messages, just as Googlebots crawl and index websites. Based on the robot’s results, Google makes money by inserting contextual advertisements into Gmail inboxes. To trade users’ privacy for money is obviously not a desirable thing, even though this might be a very lucrative business. It is natural to conclude that Code helps the Market to unlock its

⁵⁸ Danny Sullivan, “What Are Doorway Pages?” *SearchEngineWatch*, June 2, 2001 @ <http://searchenginewatch.com/webmasters/article.php/2167831>.

⁵⁹ Code swapping is also a frequently used SEO technique: it optimizes a page for top ranking, then swaps another page in its place once a top ranking is achieved. See *Fact Index*, “Spamdexing”, @ <http://www.fact-index.com/s/sp/spamdexing.html>.

⁶⁰ Trevor Haywood, “Global Networks and the Myth of Equality—Trickle Down or Trickle Away?” in Brian D. Loader (ed.), *Cyberspace Divide: Equality, Agency, and Policy in the Information Society* (London; New York: Routledge, 1998) p.22.

⁶¹ See the blog maintained by Charlie Ayers, Google’s head chef who publish what food Google guys eat daily @ <http://googlemenus.blogspot.com/>; see also Andrew Orłowski, “Google Values Its Own Privacy—How Does it Value Yours?” *The Register*, 13th April 2004 @ http://www.theregister.co.uk/2004/04/13/asymmetric_privacy/.

⁶² See Matthew Fordahl, “Google’s Stock Soars in First Day of Trading after Controversial IPO”, 20 August 2004 @ <http://www.cp.org/english/online/full/technology/040819/z081946A.html>.

tremendous potential to have commercial exploitation over almost everything. Here ethics and morality are left in a black and cold vacuum.

The third aspect of the trilemma is that the law is either too dumb or too lame to cage the rogue power of money when Code frequently conspires with Market. In the search engine world, there is a considerable time lag before the law catches up with new technologies. Early cases such as *Playboy Enterprises Inc. v. Calvin Designer Label*,⁶³ *Insituform Technologies, Inc. v. National Envirotech Group L.L.C.*,⁶⁴ and *Oppedahl & Larson v. Advanced Concepts*⁶⁵ mainly touch upon meta tag misuses. Very few cases have yet been filed relating to the use of some of the more advanced spamdexing techniques such as cloaking, code swapping, wikispamming, blogspamming and so on.

Even in some “simple” meta tag disputes, the law again demonstrated its poor understanding of technological issues. For example, in *Brookfield Communications Inc. V. West Coast Entertainment Corp*, the U.S. 9th Circuit Court of Appeals stated that “[u]sing another's trademark in one's metatags is much like posting a sign with another's trademark in front of one's store.”⁶⁶ However, it is common knowledge that meta tags are invisible to the naked eye, and just as Danny Sullivan critically observed, the meta tag “is not a ‘sign’ in the sense of that place on a store, to be read by consumers.”⁶⁷ This case partly reflected the court’s unfamiliarity with new technology, which might tarnish the authority and credibility of the law.

3. For Whom the Creativity Bell Tolls

As mentioned, case law is scarce and thin in the search engine world. By sharp contrast, in another contested sphere—digital copyright—law abounds and overspills. Even with more law, copyright bears a striking similarity to search engines: ultra-free markets and ever-draconian copyright law are both skillful at rendering Code technically impotent; i.e., both of them are capable of making Code weak and vulnerable. For example, the law has already disabled some versions of P2P networks like as Napster.com.

This section will focus upon the current copyright regime, which epitomizes the co-regulatory mechanism of Code and Law (CL Model). Lessig is quick to note one interesting fact: copyright law was born as a “shield” to protect, but has now matured into a “sword” to attack.⁶⁸ The author will also show, in this part, how the sword of copyright law is whetted on the hone of DRM and threatens to slaughter creativity

⁶³ *Playboy Enterprises, Inc. v. Calvin Designer Label*, Civil Action No. C-97-3204 (ND Cal. 1997); 985 F. Supp. 2d 1220 (1997).

⁶⁴ *Insituform Technologies, Inc. v. National Envirotech Group L.L.C.*, Civ. No. 97-2064 (E.D. La., final consent judgment entered Aug 27, 1997).

⁶⁵ *Oppedahl & Larson vs. Advanced Concepts*, No. 97 Civ. Z-1592 (D.C. Colo. 1997).

⁶⁶ *Brookfield Communications Inc. V. West Coast Entertainment Corp*, 174 F.3d 1036 (9th Cir. April 22, 1999); the ruling is also available @ <http://laws.lp.findlaw.com/9th/9856918.html>.

⁶⁷ Danny Sullivan, “Meta Tag Lawsuits”, *SearchEngineWorld.com* 21April 2004 @ <http://searchenginewatch.com/resources/article.php/2156551>.

⁶⁸ Lawrence Lessig, *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity* (New York: The Penguin Press, 2004) p. 99.

3.1 DRM and Copyright: A Shotgun Marriage of Code and Law

Legal nihilists are contemptuous about Law's effectiveness in controlling pervasive online piracy in P2P file swapping networks. They advocate that a workable alternative is to use Code—like DRM—to weave a much tighter net for the leaky digital environment.

However, it is well documented that no DRM technology is ever “100 per cent hacker-proof”⁶⁹ and to crack it is only a matter of time. Copyrighted works are supposed to be better protected by cryptographic system, but the results are far from satisfactory. For example, the Recording Industry Association of America (RIAA) once launched a Secure Digital Music Initiative (SDMI) to guard online music delivery by adopting encryption technology “operating on all platforms and adapting to any new ones.”⁷⁰ However, the project was overambitious. In 2000, the SDMI group issued an invitation to the digital community to challenge their system. Much to the shame of the RIAA, it received 447 entries claiming to have found loopholes in the SDMI system.⁷¹ It could be awkward for music industry to rely upon these unsecured systems for its “security.”

In order to redeem the botched and vulnerable Code, legislation proscribing anti-circumvention (as well as anti-trafficking) finally came into existence. After ratifying the 1996 WIPO Copyright Treaty, both Europe and the US adopted laws against hackers tinkering with Code on many occasions. Take the US 1998 Digital Millennium Copyright Act (DMCA) for example: Section 1201 outlawed the circumvention of “a technological measure that effectively controls access to a work.”⁷² Scholars, such as Pamela Samuelson, criticized the DMCA for going too far,⁷³ because it created a kind of “paracopyright”⁷⁴ that risks undercutting the “public's access to ideas”⁷⁵ and suppressing competing technologies.⁷⁶

In this scenario, the conjunctive use of Law and Code could be sweepingly repressive upon creativity as well as freedom of speech. For example, when Edward Felton, a Princeton professor, studied the SDMI system, he found several technical flaws in it. The SDMI group was reluctant to see Felton's critical research about them disseminated. Backed up by the DMCA, they wrote Felton a letter, warning him not

⁶⁹ Nick Hanbidge, “Protecting Rights Holders' Interests in the Information Society: Anti-Circumvention; Threats Post Napster and DRM” at (2001) 12 *Entertainment Law Review* 223, at 225.

⁷⁰ See Rohan Massey, “Music to Your Ears? The Ongoing Problems of Online Music Delivery”, (2001) 12 *Ent.L. R.* 96, at 97.

⁷¹ See John Lettice, “Hacker Research Team Disputes ‘Hack SDMI’ Results”, *The Register*, 9 November 2000 @ http://www.theregister.co.uk/2000/11/09/hacker_research_team_disputes_hack/.

⁷² See 17 U.S.C 1201 (2000).

⁷³ See Pamela Samuelson, “Intellectual Property and the Digital Economy: Why the Anti-Circumvention Regulations Need to Be Revised”, (1999) 14 *Berkeley Technology Law Journal* 519.

⁷⁴ Paracopyright protects rights that differ from but are associated with copyright. See Katherine C. Hall, “Deconstructing a Robotic Toy: Unauthorized Circumvention and Trafficking in Technology”, (2004) 20 *Santa Clara Computer & High Technology Journal* 411 at 415.

⁷⁵ Rod Dixon, “Breaking into Locked Rooms to Access Computer Source Code: Does the DMCA Violate a Constitutional Mandate When Technological Barriers of Access Are Applied to Software?”, (2003) 8 *Virginia Journal of Law and Technology* 2.

⁷⁶ See Dan L. Burk, “Anticircumvention Misuse”, (2003) 50 *UCLA Law Review* 1095 at 1113.

to publish criticism. Professor Felton was outraged and filed a suit against the RIAA in a New Jersey federal court. Unfortunately, the court dismissed his complaint.⁷⁷

Similar cases are not rare. In *Microsystems Software, Inc. v. Scandinavia Online*, Microsystems developed *Cyber Patrol 4*, a censorware system, which had arbitrarily blacklisted a human rights websites such as Amnesty International, a U.S. Congressman's campaign site, "Lloyd Doggett for Congress," and others. Two hackers found these problems by decompiling *Cyber Patrol 4* and, as a result, wrote *CPHack* to disable the misleading and unreliable censorware.⁷⁸ In an attempt to extinguish widespread criticism, the plaintiff strategically "purchased" *CPHack* (which was originally licensed under the GPL). Judge Edward Harrington decided that no one could use and disseminate "CPHack" because of the plaintiff's copyright.⁷⁹ Lessig vehemently disagrees with the injunction and laments: "the DMCA thus reaches more broadly than copyright law."⁸⁰

The law is harsh on hackers who develop emancipatory codes against repressive codes such as DRM. Classic cases like *Universal City Studio v. Eric Corley*⁸¹ (2600.com's dissemination of Linux-based DeCSS was banned) and *United States v. Elcom, Ltd.*⁸² (a Russian hacker was arrested by the FBI for criminal violation of the DMCA) are really disheartening, if not horrifying, to digital dissenters who demonstrate their civil disobedience against electronic orthodoxy. The cause of ethical hackers to maximize free flow of information faces serious persecution from Law. Should a scholar's studies on a technical control system be outlawed indiscriminately? Or does the entertainment industry have a secret agenda to eventually define all hackers' efforts to co-develop their code as a kind of "thoughtcrime"?

3.2 Copycrime Prevention: Imposing Copyright Levies

Content industry oligarchs are dedicated contributors to Orwell's *Newspeak Dictionary*. "Copycrime"⁸³ is their most recent contribution. *Copycrime*, in the family of *thoughtcrime*, *sexcrime*, *facecrime*, is also an all-inclusive term. *Copycrime* has the

⁷⁷ See Rohan Massey, "Anti-Copying Technology—Freedom for Speech of IPR Infringement", (2002) 13Ent.L.R. 128, at 129.

⁷⁸ See Peacefire.org, "Cyber Patrol Examined", @ http://www.peacefire.org/censorware/Cyber_Patrol/.

⁷⁹ See *Microsystems Software, Inc. v. Scandinavia Online* AB, 98 F. Supp. 2d 74 (D.Mass., 2000), aff'd, 226F. 3d 35(1st Cir., 2000).

⁸⁰ See Lessig, "Battling Censorware", *The Industry Standard*, September 8, 2001 @ <http://www.lessig.org/content/standard/0,1902,13533,00.html>.

⁸¹ *Universal City Studios, Inc. v. Corley*, 273 F.3d 429 (2nd Cir. 2001).

⁸² *United States v. Elcom Ltd.*, 203 F. Supp.2d 1111 (ND Cal. 2002).

⁸³ My coinage of the word "copycrime" is inspired by George Orwell's dystopian vision from 1984. *Newspeak*, the official language of Oceania, is a project committed to cripple critical thinking by destroying words actively.

A philologist compiling the 11th edition of *Newspeak Dictionary* explained, "... our chief job is inventing new words. But not a bit of it! We're destroying words –scores of them, hundreds of them, every day. We're cutting the language down to the bone....It's a beautiful thing, the destruction of words." See Orwell, *Nineteen Eighty-Four* (Harlow: Longman, 1983) p.43.

ambition to banish from the legal lexicon all the following words: piracy, bootlegging, private use, fair dealing, educational use, reuse, ripping, mixing, burning, reverse-engineering, technical circumvention, and derivative creation, among others.

The dystopian vision of *copycrime* is, to some extent, exaggerated. It seems that it would still take a long time before the content industry translates all concerned actual or potential infringing acts into *copycrime*. However, their persistent endeavor to crack down on *potential* “copy-criminals” is a well-documented reality.

The music industry uses an effective *copycrime* prevention technique known as copyright levies or private use levies. The levies are imposed upon digital devices (e.g. PCs, MP3 players) or blank media (e.g. CD-R/RWs) *merely capable of copying music*. They are a hidden tax in the price of the copying devices or media: a consumer is now taxed when listening to music. The logic of prepaid copyright levies can be translated into a formula: every one is a *potential* pirate and should be penalized before committing a *copycrime*. In other words, people are required to purchase the quota of being “copy-criminals.”

Most continental European countries have imposed various copyright levies, and the EU Copyright Directive (Directive 2001/29/EC) supports this practice. Article 5 (2)(b) of the Directive permits the act of private copying for non-commercial purpose by “a natural person,” on the condition that right holders “receive fair compensation.”⁸⁴ Germany is the most aggressive example and its levy system is mandated under the current German Copyright Act (section 53 *et seq.*).⁸⁵ Other countries, such as France (under *Code de la Propriété Intellectuelle*),⁸⁶ Spain (under Art. 5 of the Law on Intellectual Property),⁸⁷ Netherlands (under 1990 Amendment to the Copyright Act),⁸⁸ and Italy (under *Decreto Legislativo n.68*)⁸⁹ also have well-established private copy levy systems. More importantly, in implementing the Copyright Directive, these countries all have extended levies to digital equipment or blank digital media.

Among Common Law systems, Canada recently became the first country to systematically impose levies upon blank digital media. On 20 March 1997, the House of Commons of Canada passed *Bill C-32 (Projet De Loi C-32)*, an amendment to the

⁸⁴ Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, *Official Journal*, L 167 , 22/06/2001 p. 0010-0019, Art. 5 (2)(b).

⁸⁵ Copyright Act of the Federal Republic of Germany, 9 September 1965. The 1965 German Copyright Act was later amended for five times (i.e. 1985, 1990, 1994, 1995 and 1998).

⁸⁶ L.311-5 *Code de la Propriété Intellectuelle*.

⁸⁷ The Law on the Intellectual Property (*Texto Refundido*, January 1st, 1996 and Royal Decree 1434/1992 of 27 November); For a detailed discussion of Spanish IP law and music industry, see Javier Diaz Noci, “Copyright and Digital Media in Spain: The New Judicial Challenges”, at *Expert Meeting of Cost A20: The Future of Music , Internet and Copyright*, Amsterdam 29 Nov. 2002.

⁸⁸ Dutch Copyright Act 1912, amended on May 30th 1990.

⁸⁹ *Decreto Legislativo* 9 Aprile 2003, n. 68: *Attuazione della direttiva 2001/29/CE sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione*. (GU n. 87 del 14-4-2003-Suppl. Ordinario n.61); a link of the decree can be found at the website of FIPR @ <http://www.fipr.org/copyright/guide/italy.htm>.

Copyright Act (Canada).⁹⁰ Under the controversial 2003/2004 tariff scheme, Canada has extended a levy of up to 25 dollars (CAD) per gigabyte to MP3 players⁹¹ (See *Figure 3*). It has given rise to heated debate over the desirability and legality of the new measures. It is noteworthy that Canada has not ratified the WIPO Copyright Treaty and that it also lacks digital copyright legislation (like the DMCA). This is why Canada has become a test case for copyright levy systems in the digital age. It is also deeply worrying that the Canadian levy regime might have some indirect impact on its neighbour—the United States—forcing it into imposing a similar draconian copyright levy system.

Figure 3: The Controversial 2003/2004 Levies on Digital Media in Canada

BLANK MEDIA TYPE	2004 LEVY	2003 LEVY
CD-R and CD-RW (Non-Audio)	\$0.21	\$0.21
CD-R Audio, CD-RW Audio and MiniDisc	\$0.77	\$0.77
Audio Cassettes of 40 Minutes or More in length	\$0.29	\$0.29 per tape
Non-removable Memory Embedded in a Digital Audio Recorder (such as MP3 Players)	\$2 (up to 1 GB) \$15 (1 to 10 GB) \$25 (over 10 GB)	no levy
DVD-R and DVD-RW	proposed levy rejected	no levy
Flash Memory – Removable	proposed levy rejected	no levy

Source: Copyright Board of Canada; Canadian Private Copying Collective⁹²

In summary, Code does not replace or alienate Law. DRM actually befriends copyright levies as well as anti-circumvention law. The World Semiconductor Council (WSC) noticed a trend in the music industry: the “ever-increasing levy rates imposed on a broadening range of equipment and media and the ‘increasing use of DRM’ would actually run ‘in parallel.’”⁹³ The co-existence of DRM and copyright levies could be a double burden upon consumers that could be too heavy to be discharged. Though the levies are imposed on the hardware/blank media suppliers, the cost are then passed on to consumers. The danger of “double taxation” exists when the technical access controls and the copyright levy are both in work. Under a blanket levy, law-abiding consumers are penalized in conjunction with copyright infringers, so they may feel discouraged to respect copyright. This is a vicious circle: the harsher

⁹⁰ Bill C-32, as passed by the House of Commons of Canada, an Act to amend the Copyright Act, March 20, 1997.

⁹¹ Copyright Board of Canada (CBC), *Fact Sheet--Copyright Board's Private Copying 2003-2004 Decision*, December 12, 2003, @ <http://www.cb-cda.gc.ca/news/c20032004fs-e.html>.

⁹² *Ibid.*

⁹³ WSC, *Draft WSC White Paper on Copyright Levies—The Dangers of Copyright Levies in a Digital Environment*, 2003, p.2 @ http://www.eeca.org/pdf/11b_levies_paper_final.pdf.

the crackdown on consumers, the less likely they are to be positive about copyright law.

4. Conservation of the Public Domain

Both the CM model and CL model of regulation are *control* theories, dedicated to the definition of constraints over human behaviors by using Code. In short, they are struggling to work out what you should *not* do. It is also instructive to note that the CM Model and the CL Model represent two institutional devices for information policy: (1) the CM Model is self-regulation under market orientation, and (2) the CL Model invites government intervention.

Yochai Benkler is not satisfied with the above two mainstream regulation approaches. He advocates that apart from direct government intervention and privatization there should be a third way: he calls it a “commons approach.”⁹⁴

Policy makers, who are only obsessed with the market-versus-government debates, have long ignored the idea of a commons. According to Benkler, a commons-based information policy is designed to “increase the degree of decentralization that can be sustained within the institutional constraints our society imposed on information production and exchange.”⁹⁵ In contrast to the CM and CL models, the commons thesis constructs a normative order—it helps to realize the emancipatory power of Code to cultivate and nurture the public domain in cyberspace, regardless of being governmental or non-governmental.

4.1 Code Should Build Upon A Normative Order

The commons approach signals a paradigm-shift. It is against the economic determinism of the CM Model and the legal rigidity of the CL Model in an established order, whose old mentality is unable to accommodate the changes brought by rapidly developing information technologies. An ultra-free market system lashes human beings into insatiable animals greedy for monetary profits. A rigid legal system deteriorates itself into as “a conservative force”⁹⁶ responding only to the needs of politico-economic oligarchies. In short, the economic-spurring mechanism and legal inertia do not hold much normative value for cyberspace.

Code should be built upon on a normative order by enabling and connecting people from all backgrounds on a nearly frictionless medium. The normative order is not a collection of empty moral rules. It would materialize itself into practical guides for policy makers in concrete situations. Scottish legal philosopher Neil MacCormick once stressed the practicalities of a “normative order” when we make choices, easy or hard:

Normative order is a kind of ideal order. At any given time we may form a view of the world as we think it is, including the set of ongoing human actions and intentions for action. We may set

⁹⁴ See, generally, Yochai Benkler, *supra* note 4.

⁹⁵ *Ibid.* at 27.

⁹⁶ Kathy Bowery, “Ethical Boundaries and Internet Culture”, in Lionel Bently and Spyros M. Mantias, *Intellectual Property and Ethics* (eds.) (London: Sweet & Maxwell, 1998) p.6.

*against that a view of the world as it could be or could become, leaving out certain of the actions, leaving some actual intentions abandoned or unfulfilled, while other actions take place instead of those left out, and other intentions are fostered and brought to fulfillment.*⁹⁷

Code is in flux and is not self-defined. Code is, as Richard A. Spinello said, always “in the hands of the state or individuals because of its *malleability* and *obscurity*.”⁹⁸ Wiebe E. Bijker, among others, developed the powerful theory of Social Construction of Technology (SCOT), which elegantly and forcefully explains how technologies are shaped by heterogeneous social forces.⁹⁹ Using a conceptual shortcut, the author proposes that technologies defining Internet infrastructure (i.e. Code) are mainly determined by three social forces—Market, Law and Norm (See *Figure 4*). These three forces either collaborate or compete.

The examples in the previous discussion have demonstrated that both Law and Market are able to manipulate the algorithmic blindness of Code, making cyberspace a cash machine for the powerful and the rich.

As a precious alternative, Norm should play a major role in leading Code. Norm provides a guide, showing the direction at which Code should turn. When Law and Market dispute and compete for primacy, Code is at a crossroads. Then the Norm steps in and calls Code to have a balanced view. The Norm prescribes a commons approach, aiming at an unencumbered public domain. Brett Frischmann believed that “the Internet infrastructure is a social construction that is not inherently a public good or a private good.”¹⁰⁰ In this light, the public domain is designed and manufactured. Norms of cyberspace are to enlarge and preserve the public domain. In other words, the public domain is a norm-friendly platform for people—to the largest extent—to share information upon which they could draw correct decisions. Philosophically speaking, the Code of Internet should be designed to facilitate what Jürgen Habermas called *öffentliches Raisonement*, i.e., the people’s public use of their reason.¹⁰¹

⁹⁷ Neil MacCormick, *Questioning Sovereignty: Law, State and Nation in the European Commonwealth* (Oxford: Oxford University Press, 1999), pp.3-4.

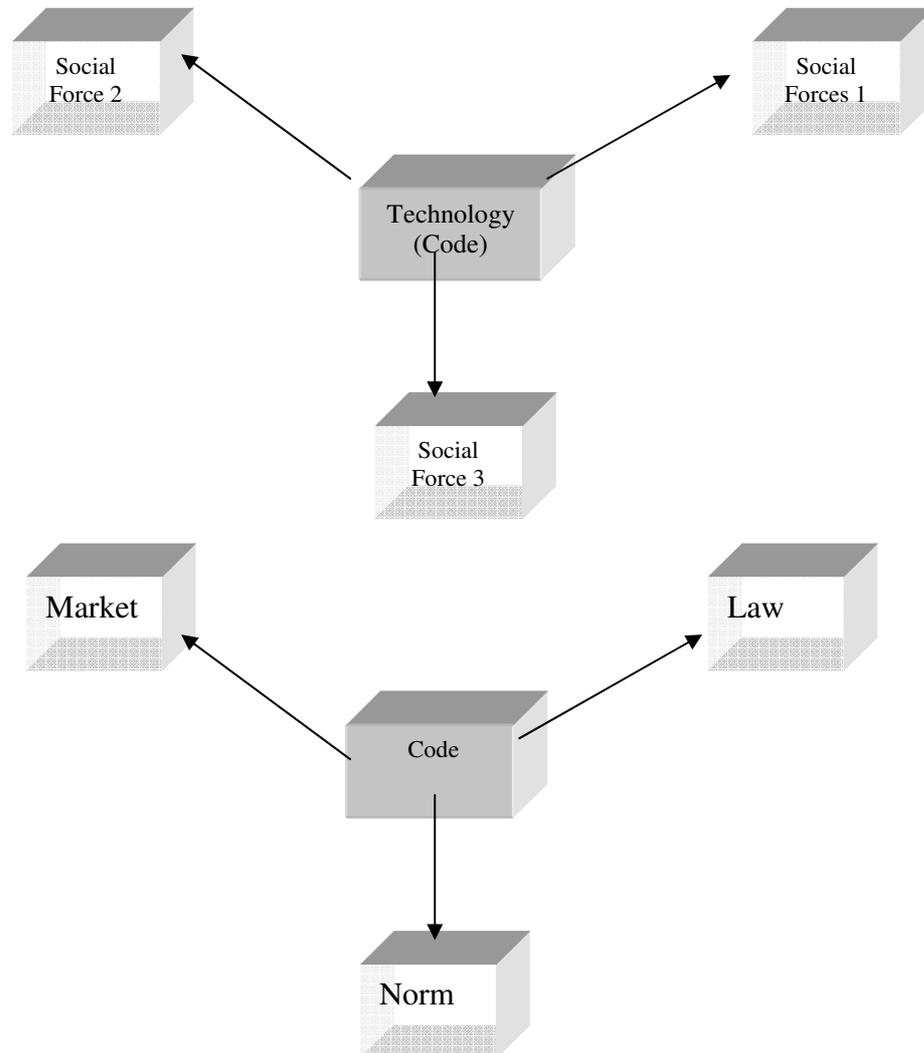
⁹⁸ See Richard A. Spinello, *CyberEthics: Morality and Law in Cyberspace* 2nd ed. (Boston: Jones and Bartlett Publishers ; 2003) p.50.

⁹⁹ See, generally, Wiebe E .Bijker, *Of Bicycles, Bakelites, and Bulbs : Toward a Theory of Sociotechnical Change* (Cambridge, Mass.; London : MIT Press, 1995).

¹⁰⁰ Brett Frischmann, “Privatization and Commercialization of the Internet Infrastructure: Rethinking Market Intervention into Government and Government Intervention into the Market”, (2001) 2 *Columbia Science & Technology Law Review* 1, p.68.

¹⁰¹ Jürgen Habermas, *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*, trans. by Thomas Burger, with the assistance of Frederick Lawrence (Cambridge: Polity Press, 1989) p.27.

Figure 4: Code is Malleable: Social Construction of Technology/Code



Source: adapted from Wiebe E. Bijker¹⁰²; Richard A. Spinello¹⁰³

4.2 “Welcome to the Jungle!”: Commons and WikiWiki

Niklas Luhmann warned that technology was losing the quality of justifying itself, because rapidly growing technology had deteriorated into a mere “installation”:

*The form of technology completely loses its quality as a form of rationality...Technology is a[n] ... installation. Thus the increasing use of technology in modern society and the description of the world on the model of technology... does not imply a judgment on the rationality of society.*¹⁰⁴

¹⁰² The figure is adapted from Wiebe E. Bijker, supra note 103, p.47.

¹⁰³ See Spinello, supra note 102.

¹⁰⁴ Niklas Luhmann, *Risk: A Sociological Theory*, trans. by Rhodes Barrett (Berlin: Walter de Gruyter, 1993), p.88.

After the technology is “installed,” it is not then responsible for whatever consequences it brings. In 1996, the European Commission issued a monumental Green Paper concerning living and working in information society. The Green Paper was highly concerned about the uncertainties that might offset the *benefits* brought by IT development.¹⁰⁵

The Internet commons is a gigantic pool, free from ownership and control. However, the vastness of the pool also betrays a sense of chaos. Useful or relevant information cannot always be retrieved effectively, and people are suffering from more and more obtrusive, unwanted, and distracting junk information. Historian Theodore Roszak worried that we are facing “the cult of information,”¹⁰⁶ in which the overabundance of information generates a surprisingly wide range of negative effects. In this light, the commons approach should not only make an ocean of information available on the Web, but also address the issues as to how to optimally organize and exploit the commons information.

Many people are skeptical about the viability of the commons model, and their suspicions are legitimate. How would a commons—by providing large quantity of free information—be able to co-exist, if not compete with, the established capitalist/economic model? The answer is that the Internet commons operates largely on the same principles as gift economies.¹⁰⁷ Unlike CM and CL models, which rely heavily on property rights and economic return, the commons model breeds a sense of community and togetherness. The underlying norm is that the more you contribute, the more diverse the pool will be, and the more return you can get from it.

A striking example of a commons model project is Wikipedia—a free online encyclopedia. Wikipedia’s success demonstrably shows how *informational altruism* outperforms “laissez-faire selfishness.”¹⁰⁸

Again, the Norm precedes the Code. Stallman wrote his prophetic “The Free Universal Encyclopedia and Learning Resource” before Wikipedia was technically possible. He prophesized that “teachers and students at many colleges around the world will join in writing contributions to the free encyclopedia.”¹⁰⁹ The *wiki* (or *WikiWiki*) technology, which allows universal access and universal editing through a browser, let Stallman’s dream come true (*Wikiwiki* is originally a Hawaiian phrase, meaning “Quick Quick”).

It is interesting to note that Wikipedia is built on the ruins of Nupedia (www.nupedia.com). Nupedia was an online commercial encyclopedia that soon

¹⁰⁵ European Commission, Directorate General V, *Green Paper—Living an Working in the Information Society: People First*, 31 December 1996, Belgium, Draft 22/07/96, Final COM(96) 389.

¹⁰⁶ See, generally, Theodore Roszak, *The Cult of Information: A Neo-Luddite Treatise on High Tech, Artificial Intelligence, and the True Art of Thinking* (2nd ed.) (Berkeley, Calif.; London: University of California Press, 1994).

¹⁰⁷ For a detailed discussion on gift economy, See Richard Barbrook, “The Hi-Tech Gift Economy”, *First Monday*, Volume 3 Number 12, 7th December 1998, @ http://www.firstmonday.dk/issues/issue3_12/barbrook/.

¹⁰⁸ Richard Stallman, “Why Software Should Be Free”, 24 April 1992, *Free Software Foundation* @ <http://www.gnu.org/philosophy/shouldbefree.html>.

¹⁰⁹ Richard Stallman, “The Free Universal Encyclopedia and Learning Resource”, *Free Software Foundation* @ <http://www.gnu.org/encyclopedia/free-encyclopedia.html>.

failed. The Nupedia team spent more than \$250,000 USD on an editor and to develop software.¹¹⁰ Shamefully, it was only left with 23 finished articles and 68 incomplete entries despite the large initial investment.¹¹¹ Then the founders of Nupedia turned to Wikipedia. The new project is open and free.

Wikipedia is pushed by Norm, not Market. It revitalizes the spirit of the 18th century *Encyclopédie* by using 21st century “peer production”¹¹² technology. The English version of Wikipedia currently has 337,911 articles (calculated on 1 September 2004), compared to 270 “articles” (all pages counted) on 25 January 2001.¹¹³ The articles/entries are peer-reviewed and maintain a high quality. Admittedly, wiki vandalism (i.e. “bad-faith addition, deletion or change to content” in editing Wikipedia) is not rare.¹¹⁴ However, according to an IBM study (done by IBM’s Collaborative User Experience Research Group), most vandalism can be fixed within a few minutes by peer writers¹¹⁵ (it was reported that median time for recovery was 5 minutes¹¹⁶). Moreover, Wikipedia is a live book, abreast of new socio-technological development. It is always much quicker to incorporate new words than commercial encyclopedias. For example, neither [Britannica Online](#) nor MSN Encarta has entries for “copyleft” or “autopoiesis”, while Wikipedia contains very detailed information about them.

Wikipedians are not legal anarchists. Instead they are sensitive to copyright. Wikipedia adopts *GNU Free Documentation License V1.2*,¹¹⁷ which made it the first copylefted encyclopedia. This license protects the collaborative work from various misuses and assures it is disseminated most effectively.

The biggest headache wiki projects recently suffer is wikispamming. Wikispammers flood wikis with links to their commercial websites to increase their ranking in search engines (see 2.3 above). The *openness* of a wiki is therefore a double-edged sword: it allows a wide range of people to join a collaborative project, while it also leaves wikis vulnerable to attacks from spamdexers. Michael Froomkin suggests that electronic vandalism, like wikispamming, is forcing wiki communities to find ways to police themselves.¹¹⁸ Interestingly, Froomkin’s own wiki project—*Copyright Experiences*—

¹¹⁰ *Far Eastern Economic Review*, “Encyclopin-off”, 19 February 2004 p.39.

¹¹¹ Wikipedia, “Nupedia”, @ <http://en.wikipedia.org/wiki/Nupedia>.

¹¹² See, generally, Yochai Benkler, “Coase’s Penguin, or, Linux and the Nature of the Firm”, (2002) 112 *Yale Law Journal*, 369.

¹¹³ See Wikipedia, “Wikipedia: Size of Wikipedia”, @ http://en.wikipedia.org/wiki/Wikipedia:Size_of_Wikipedia.

¹¹⁴ See Wikipedia, “Wikipedia: Vandalism”, @ <http://en.wikipedia.org/wiki/Wikipedia:Vandalism>.

¹¹⁵ The detailed information for the research (about its methods and results), see IBM’s Collaborative User Experience Research Group, “History Flow—visualizing dynamic, evolving documents and the interactions of multiple collaborating authors: a preliminary report”, whose project homepage @ <http://researchweb.watson.ibm.com/history/index.htm>; the results are available @ <http://researchweb.watson.ibm.com/history/results.htm>.

¹¹⁶ See JOI ITO, “Wikipedia Heals in 5 Minutes” @ http://joi.ito.com/archives/2004/09/07/wikipedia_heals_in_5_minutes.html.

¹¹⁷ FSF, *GNU Free Documentation License V 1.2*, November 2002 @ <http://www.gnu.org/copyleft/fdl.html>.

¹¹⁸ A. Michael Froomkin, “Habermas @ Discourse.Net—Toward a Critical Theory of Cyberspace” (2003).

is also wikispammed by a host of malicious websites these days. Froomkin implemented a primitive way of fighting spamming by blocking users from some IP addresses:

*8/17/04 - Banned all links to *.com.cn (crude way to block spam, but it's getting out of control). Banned several IPs too.*

*7/16/04 - Banned all access from *.netvision.net.il (apologies to legit users from there, if any).*

7/05/04 - Removed WikiSpam from a casino that had infected the site on many pages.¹¹⁹

However, IP address banning is not very effective to stop spamming. It is easy for spammers to use a new IP address, after the original one is blocked. The blanket banning of (*.com.cn) is rough justice: it also exclude all *legitimate* wiki users affiliated to these domain names. Wikispamming is a perfect example of Garrett Hardin's tragedy of commons.¹²⁰ The commons is misused or abused by a few to promote their private interest.

Again, we need Norm to purify the wiki communities by restrict wikispammer's selfish commercial acts. There is a morality for the Web. Wiki users are not nihilists; and they want wikispammers to understand that their efforts to externalize their advertising costs to the commons will only result in re-internalizing those costs in the end.¹²¹ It is encouraging to see that two graduate students voluntarily set up an ethical website—*Chongqed.org*—to retaliate against wiki and blog spamming. Their way of fighting is simple and clever. They advise people to replace spammers' links with *Chongqed.org* links and thus frustrate spammers' commercial motivation by hurting their PageRank:

We encourage others to set links to chongqed.org using the same keywords that the spammers used while vandalizing wikis and blogs. This gets us a good page rank for the spammy keywords and this gives us the chance to inform others about wiki spam, wiki spammers, spamvertized web sites, etc.¹²²

The heavier the spamming, the more spammers' links will be replaced. Finally, *Chongqed.org* would manage to outperform these spammy sites in search engine results and make spamming ineffective.

116 *Harv. L. Rev.* 749 at 862.

¹¹⁹ See *Copyright Experience Wiki*, "ChangeLog" @ <http://www.discourse.net/commons/wiki.pl?ChangeLog> (accessed on 5 September 2004).

¹²⁰ See Garrett Hardin, "The Tragedy of the Commons", (1968) *Science*, Vol. 162, Issue 3859, 1243-1248.

¹²¹ Ira S. Nathenson, "Internet Infoglut and Invisible Ink; Spamdexing Search Engines with Meta Tags" (1998) *Harvard Journal of Law & Technology* 43 at 131.

¹²² See Chongqed.org, "Fighting Wiki Spam" @ <http://chongqed.org/fightback.html>.

5. Conclusion: Code Prescribes a Digital “Shock Therapy” for Law’s Illness

Freedom and control go hand in hand. Just as French philosopher Michel Foucault said, “power and freedom’s refusal to submit cannot therefore be separated.”¹²³ Where there is digital freedom there is control and constraint. The sound understanding of the dialectics of freedom and control is a key to the role that the Code plays in cyberspace.

Code either regulates or emancipates. This depends on what kind of interplay Code has with the other three regulatory modalities, i.e. Market, Law and Norm. Code is not self-determined and is subjected to social molding. This results in a three-pronged information policy strategy (See again *Figure 1*):

1) The CM Model (self-regulation model):

Code is led by Market when there are few laws. We must be alert to the fact that the Internet originally was only for science and defense research. It started from Advanced Research Projects Agency Network (ARPANET) and was developed by the U.S. Department of Defense, who were aiming to create a network architecture that “can not be controlled from any center.”¹²⁴

The Internet sinned against the non-commercial and sharing norm when its infrastructure was continuously privatized in 1990s.¹²⁵ It was no longer the Garden of Eden, but full of evil spirits for money. The Law and Economics Movement¹²⁶ initiated by the so-called Old Chicago School tends to interpret every change in terms of Market. Posner and Landes contend that “economics can reduce a mind-boggling complex of statutes, amendments, and judicial decisions to coherency.”¹²⁷ The economic approach is useful but does not go unchallenged. Neil Netanel is skeptical about the market-oriented regulation theory:

*I question the democratic nature of the market-based decision-making process. That process is heavily biased towards those with the financial resources, generally acquired outside of cyberspace, to expend on convincing others (perhaps by words, perhaps by bribes, perhaps by threatening to block their sites and servers) to join one’s side.*¹²⁸

¹²³ Michel Foucault, “The Subject and Power”, *Afterword* in Hubert L. Dreyfus and Paul Rabinow, *Foucault: Beyond Structuralism and Hermeneutics* (Brighton: Harvester, 1982) p.221.

¹²⁴ See Manuel Castells, *The Rise of the Network Society* (Malden, Mass ; Oxford : Blackwell, 1996) p.6.

¹²⁵ See Granville Williams, “Selling Off Cyberspace” in Stephen Law (ed.), *Access Denied in the Information Age* (Basingstoke: Palgrave, 2001) p.190.

¹²⁶ See Richard Nobles, “Economic Analysis of Law”, in James Penner, David Schiff and Richard Nobles (eds.), *Introduction to Jurisprudence and Legal Theory: Commentary and Materials* (London : Butterworths, 2002) pp.855-858.

¹²⁷ William M. Landes, Richard A. Posner, *The Economic Structure of Intellectual Property Law* (Cambridge, Mass.; London : Belknap Press of Harvard University Press, 2002) p.10.

¹²⁸ Neil Weinstock Netanel, “Cyberspace Self-Governance: A Skeptical View from Liberal Democratic Theory” (2000) 88 *California Law Review* 395 at 472.

In the search engine market, we have seen how search engines and parasitic SEO businesses conspire to distort “natural” search results when lured by profits. Code could be easily misled by market forces. Moreover, the CM model suffers from serious democratic deficits and unsavory social Darwinism, because it implies “the survival of the fittest... with little regard for the rights of the technologically weak.”¹²⁹

2) The CL Model (government intervention model):

Legislation does not always reflect public opinions. In the high-profile case of *Eldred v. Ashcroft*, Justice Breyer, in dissent, commented that “the statute might benefit the private financial interests of corporation or heirs who own existing copyrights.”¹³⁰

Content industries treat P2P technology as their arch-enemy. They believe the Code of new technology to be an erosion of the established copyright system. However, they cannot understand the fact that new technologies can serve as a shock therapy to cure the legal illness. P2P networks are a “creative destruction”¹³¹ to an over-broad copyright regime that is good at suffocating creativity and competition. Since Law is reactionary and technology is forward-looking, the balance should be reached through a normative order.

3) The Commons Model:

The commons approach weakens the antagonism between self-regulation and government regulation. Norm should lead, since Code is easily maneuvered by the Market or by Law for private interests.

The Norm derives from the value-laden architecture of the Internet, whose openness enables “a virtuous cycle where members of the Internet community continued to improve upon its basic architecture by adding new functionalities that were placed in the public domain.”¹³² A causal link should also be established between the Internet and democratization.¹³³

Norm is to conserve a public domain in which all informational resources are freely accessible to everyone. The public domain is an eco-system. James Boyle argues that, like the idea of a natural environment, the commons must also be “invented” before it is saved.¹³⁴ Not surprisingly, commercial contamination of the Internet has already made information pollution a persistent problem. The disturbing problem of

¹²⁹ Henning Wiese, “The Justification of the Copyright System in the Digital Age” (2002) 24 *EIPR* 387 at 390.

¹³⁰ *Eldred v. Ashcroft*, 123 S.ct.769 (2003).

¹³¹ See Raymond Shih Ray Ku, “Consumers and Creative Destruction: Fair Use Beyond Market Failure”, 18 *Berkeley Technology Law Journal* 539.

¹³² Philip J. Weiser, “Internet Governance, Standard Setting, and Self-Regulation” (2001) 28 *Northern Kentucky University Law Review* 822 at 826.

¹³³ Jane Strachan, “The Internet of Tomorrow: The New-Old Communications Tool of Control”, (2004) 26 *EIPR* 123 at 123.

¹³⁴ James Boyle, “The Second Enclosure Movement and the Construction of the Public Domain” (2003) 66 *Law and Contemporary Problems* 33, p.52.

information dumping (e.g. email spamming, spamdexing, etc.) is often induced by the Market and is hardly fixed by Law. A strong normative order needs to be established to contain them.

In summary, the Code should be led by Norm to emancipate and realize its unlimited potential while acting within an ethical boundary. The Code of new ITs foretells the Age of Electronic Renaissance. Code, following the Norm, also enriches and substantiates the *lex informatica*¹³⁵ and it evokes the most meaningful discourse in the digital age: the idea of commons. John Donne, a 17th century religious poet, once wrote “no man is an island, entire of itself.”¹³⁶ In fact, like undivided humanity, commons belongs to no one because commons belongs to everyone: then never ask for whom the bells tolls; “it tolls for thee!”

¹³⁵ See Joel R. Reidenberg, “Lex Informatica: The Formulation of Information Policy Rules Through Technology”, (1998) 76 *Texas Law Review* 553, at 579.

¹³⁶ John Donne, “Meditation XVII” (1624) also available @ <http://www.geocities.com/Yosemite/Trails/2237/lit/Meditation17.html>.