Predicting the evolution of digital rights, digital objects, and digital rights management languages.

Jonathan Schull, Associate Professor, Information Technology, Rochester Institute of Technology

***Abstract*— As a sometime biological psychologist and sometime DRM pioneer, I suggest that biological principles are at least as important as technological principles in anticipating future developments in the field of rights management, and requirements for digital rights languages. Among those possible developments are (1) increases in the virtuality and virality of rights-managed objects, of distribution systems, and of payment systems, (2) systems for tracking the copying and redistribution of digital documents, (3) application of digital rights to data derived from document tracking, (4) attribution of those rights to the individuals who do the re-distributing, (5) development of rights management systems for the aggregation, protection, anonymization, and monetization of personal information, (6) rights-managed digital objects whose content changes spontaneously as a function of normal use, and (7) digital objects that adapt through a natural selection-like process of mutation, recombination and differential reproduction.**

Such ideas pose interesting challenges for rights management languages.

***Index Terms*—Copyright Protection, Rights Management, Superdistribution, Natural Selection**

I. Introduction

ANY rights management language that hopes to keep pace with "facts on the ground" must be extensible to rights management practices that are uncommon, but predictable, today. We can better design a digital rights language for the future if we can anticipate the changes and change processes we will have to accommodate.

As a sometime biological psychologist and sometime DRM pioneer [1,2,3], I believe that biological principles are at least as important as technological principles in anticipating future developments in the field of rights management.

Today's digital rights management situation represents the convergence of two historical trends: virtualization and biologization. Of the two, biologization is the least discussed, let alone well understood. But its implications are most fundamental for digital rights management and for the transformation of the information economy.

Manuscript received June 13, 2005.

Author is with the Interactive Media Group in the Information Technology Department at the Rochester Institute of Technology, Rochester New York 14607. Phone: 585-738-6696. email schull@digitalgoods.com

Here's what I mean. Traditional economies are based upon the delivery of valuable “things” (products and services) in exchange for receipt of valuable “things” (including money). However, starting at the dawn of civilization, value came to be represented first by tokens, then by coin, then symbolically in money, and then virtually in disembodied bits. That's virtualization: symbolic representation with more and more impact with less and less mass and energy.

At the dawn of life, value was embodied first in analog form in the biological processes of single celled organisms, and then symbolically in digital form by RNA and DNA. That too is virtualization. But life also teaches us that when things are virtualized, reproduction becomes easier, and biological dynamics of reproduction and evolution arise. Virtualization is a step on the road to biologization.

Today, information products are being virtualized. To fully understand alternative rights management options, it may be helpful to look closely at a spectacularly successful economy based not on state-sanctioned currencies, but upon unregulated reproduction, competition, and innovation. That economy is all around. It is the world of biology.

My goal in this paper will be to provide a broad historical, if idiosyncratic perspective, on the past and present evolution of digital objects and rights management systems. Needless to say, these ideas are offered as useful speculations, not confident predictions, about the future.

II. Behavioral engineering and Digital Goods

In the early 1990s I was a biological psychologist and amateur programmer interested in the co-evolution of biological, social and informational ecologies [4,5]. I studied animal behavior, and had created some useful software for analyzing my data. I wanted to distribute this software, to be compensated for my work, and to take advantage of the then-emerging virtualization of software products by distributing and selling my software over the Internet. I wanted my software to reproduce, like a positive virus, so that users would "infect" their friends by making and sending copies. In those days I was literally studying and observing paramecia as they swam around, reproduced, and proliferated; I had those images in my mind.

The shareware concept had been around for almost a decade, [6] but I knew that my customers, like me, were unlikely to assemble a check, an envelope, and a stamp if their only

reward was the delivery, weeks later, of a now-redundant diskette or a postcard-of -thanks. As a behaviorist, I knew that contingencies of reinforcement dictate that meaningful and relevant rewards should be delivered within a fraction of a second of the behaviors they are intended to encourage.

The instant reward I could use was obvious—increased access to the most valuable features of my product. But it was less obvious how I could ensure that the product would be purchased again (and again) each time it was redistributed. Before we consider that puzzle, I'd like to revisit the twin issues of virtualization and biologization. Because it turns out that my solution to this practical problem also led me to rethink my understanding of the information economy.

III. VIRTUALIZATION AND ECONOMICS

Even with shareware, money usually changes hands under the consensual delusion (or user-interface metaphor) that information products are things-- "goods"—and that publishing is a business in which manufactured things (like books) are traded for things (like gold doubloons) owned by the purchaser. The irony, of course, is that what consumers "hand" over these days typically cannot be "handled"--it is symbols (digitally encoded, perhaps in plastic credit), which give the "bearer" (who "bears" nothing) the right to control the disposition of other symbols in the future. And what consumers get back from publishers is less and less likely to be physical as well: software and music, books and movies are all moving into a realm in which delivery and consumption is the symbolically-controlled execution of virtual operations by virtual machines in virtual places "on the web" or "in the bank".

As these examples show, money was virtualized long before other forms of intellectual property. And as money became virtualized it became more and more copyable. Today it takes the constant vigilance and full force of the most powerful political and military forces in the history of mankind--governments, businesses, and the police forces that back them—to prevent money from being copied by unauthorized parties. Thus, copy-prevention is a time-honored solution to the fact that virtual goods are copyable goods.

However, while copy prevention may well be necessary to preserve the integrity of our monetary system and civilization, as we know it, it may well be counter-productive when it comes to other virtual value-objects. The thing-based transaction-metaphor adopted by commercial publishing may have outlived its usefulness.

IV. VIRTUALIZATION, BIOLOGIZATION AND THE INFORMATION ECONOMY

"Publishing" actually has two very different meanings and histories. The “thing-based manufacturing metaphor” can be said to have started with Gutenberg: books are manufactured, and exchanged for “cash on the barrelhead”. But there is also a much-older idea-based information dissemination activity called “publishing” that has been practiced non-commercially for millennia by authors, scholars, pamphleteers, theologians, by flowers (which disseminate vast amounts of genetic information and arrange to have it distributed, at little cost, on the wings of the wind.)

The essential “product” in this case, is information. And information is not a thing. It is a process by which patterns "in-form"--impress themselves upon--things. Furthermore, as we have noted, because these patterns are only loosely coupled to the media they inform, they reproduce, they spread, and they evolve. They don’t just move from place to place like traditional “things”.

To make a long story short, patterns that reproduce, spread and evolve originated in the primal soup 3-4 billion years ago, they spread into (and helped create) protocells, RNA, DNA, and organisms that make their living by in-forming their environment. Approximately 1 billion years ago, propagating patterns branched out to a new media-- animal nervous systems--that allowed them to reproduce, first via learning, then via spoken patterns of sound, then via written patterns of ink on paper, and just in the last century, as patterns of electrons in yet another culture-medium that is now known as the global internet. [7,8]

Thus, over the last century the remarkable dynamics and “technology” of biology have come to be understood. My claim is that digital rights practitioners need to recognize that those dynamics and emerging analogous technologies are an increasingly fundamental part of their own discipline.

V. NATURE'S PUBLISHING ECONOMY

The "economy of nature" depends relatively little on the principles of thing-based manufacturing economies. Plants and animals do sometimes organize reciprocal resource exchange relationships, but the resources that are exchanged are services (including reproductive services) as often than as they are things. Here's how this observation applied to my own work, and the concept of superdistribution.

You will recall that I wanted my users to copy and redistribute my software, and I wanted to be able to reward those who decided to purchase it by giving them instant access to the product's advanced features. I imagined a happy purchaser passing a copy on to a friend with a recommendation. When the friend executed the program, she would have limited access to the advanced features her friend had purchased, until she committed to a purchase. The moment she made a payment the product would provide full access. However she passed copies on to her friends the copies needed to revert to “demo” mode. Thus, I needed a lock that would respond to a combination of code plus context.

I'm sure there were other ways of getting to the right answer, but my inspiration was biology. Biological functions are embodied not in genes nor in the environment, but in the dynamic interaction of genes (code) and the environment. Change either genes or environment, and function (skin color, say) may change.

My code was not going to change; it was going to be copied perfectly (and, I hoped, often). But the environment of one user would be different from the environment of another user. So I could have my code behave differently when it detected

that it had been moved from the environment of a purchaser to the environment of a non-purchaser.

(In most systems, including mine, the environment that the software responds to is the user's computers. But in the patent I eventually wrote, and in the future, the enabling environment should be the user herself. After all, it is she who purchases the service. Rights management languages are going to have to accommodate the vagaries and constraints of biometric systems. *Can matters of biology and individuality be expressed in ODRL?)*

So here is how I ended up vending my animal behavior software. When the program started up, it profiled the user's computer, made a list of relatively stable but idiosyncratic features, added up all the ASCII values of the characters in that list (literally!) to produce a large number, used that computer "fingerprint" as the seed to a random number generator, and generated a many-digit magic "password". The program then looked for that magic number in a "password file" on the users hard drive, and if the right number could be found in that password file, it functioned in "professional mode"; if not, it functioned in "demo mode" and encouraged the user to by his own password. The nice thing about this arrangement was that even if the password file was copied and redistributed along with the software, the program would still come up in "demo mode" because the magic password for one user's machine was not valid for another user's machine.

Now, the only person who knew how to generate passwords was me. When a customer decided to purchase, she called a software vendor (by phone), he took payment (by credit card or purchase order) and wrote down the fingerprint, and he called me (by phone). I would get calls (sometimes while delivering lectures on cultural transmission and gene environment interaction) and speak the password to the vendor who would later speak it to the customer who would later type it into her password file.

After a year of this, I realized (1) this was working (2) the idea was potentially more significant (even as biology!) than the animal behavior I was trying to analyze (3) that it could be applied to software products other than mine (4) that software was a service (even though I occasionally referred to my business as a random number manufacturing and vending facility) (5) a password vending service was a good job for a computer--running from class to phone to computer to phone and back to class was silly. So the patent I wrote [1] and the business I started was (SoftLock Services *aka* DigitalGoods.com) was based on the idea of a software toolkit that could accommodate multiple authors, multiple products and multiple features, all coupling these product to a password vending system that took payments, delivered passwords, and distributed funds to software developers (and us).

VI. The history of "Superdistribution"

To my knowledge this is the earliest example of software-only "superdistribution". The term itself was invented some years earlier by a Japanese computer scientist named Ryoichi Mori who defined it as an "approach to distributing software in which software is made available freely and without restriction but is protected from modifications and modes of usage not authorized by its vendor"[9] But in fact, Mori's own system presumed the existence of special tamper-proof hardware, as did Brad Cox who popularized the concept and emphasized usage-metering in book and magazine publications around 1994[10,11] The concept was further popularized, and arguably co-opted, by Intertrust's founder Vincent Shear.[c.f. 12]

My impression is that most people think of "superdistribution" as a software-only process, like what I implemented. But in any case a software-only process is certainly more virtual and more viral than one that requires the distribution of special hardware

It's worth noting, however, that today's superdistribution concepts can be taken still further. Superdistribution could be more virtual. We don't have to assume an "earthbound" payment processing system run by a credit card processing system and linked to the banking network. With peer to peer architectures and web services, its possible to imagine a system in which software services or non-monetary information assets were the only "coin of the realm," with transactions being remunerated not with money but with scrip, redeemable for services or information assets. While some of these services would presumably have to be redeemable somewhere, somehow, for something of "nutritional" or "reproductive" value, our concept of payments as well as our products can and will go ever more virtual. *It's not clear to me whether ODRL can currently accommodate non-financial remuneration.*

A well-worked out example of non-monetary currency is "whuffie," as described in digital rights activist Cory Doctorow's science fiction novel, Down and Out in the Magic Kingdom [13] which depicts a world in which "whuffie" an constantly updated measure of reputation that motivates people to do useful and creative things. Anything is available to you if you have good whuffie, and those who make those goods available gain whuffie indirectly. But if you make a lot of enemies, your whuffie plummets. It's a good read, and except for the fact that the whuffie market is mediated by internet-connected brain implants, this futuristic scenario is actually hundreds of millions of years old: among many social mammals mating opportunities and access to environmental resources often based upon hard-earned social status.

A less outlandish example of non-monetary currencies arises when we consider compensating users for virally superdistributing content. Consumers who recommend and distribute products to their friends are providing marketing, distribution, sales, and technical support services to their recipients. Why should they not be compensated? And if we are going to compensate them, why not compensate them with something that we can "manufacture" at no cost—the right to consume other digital products?

Rights management languages will therefore face new challenges as the virtuality and virality of superdistribution arrangements increases. *Can ODRL specify compensation rights for people who redistribute but do not modify rights-managed content, and can it specify alternative currencies?*

(Incidentally during the "Great Ebook Boom of March, 2000", when Stephen King's published his ebook "Riding the

Bullet [14], I tried to determine how much redistribution was actually happening. To my surprise and dismay, there was relatively little. A survey suggested the reason--many of our customers told us they thought that that "wasn't allowed", even though our marketing materials explicitly encouraged them to pass copies to their friends. So one reason we were interested in compensating redistributors was to create some pro-copying propaganda to counter industry brainwashing that implies, with misleading simplicity, that copying violates copyrights.)

VII. Tracking information flows?

In order to compensate users for redistributing our products, we would need a good way of tracking redistribution. As a would-be "information ecologist" this was of great interest to me for other reasons as well.

First, I think that tracking the flow of digital objects and activities is a huge scientific opportunity. A field biologist once told me that hydrologists sometimes map Biscayne Bay in Florida by dropping thousands of oranges into the water, and taking aerial photos a day later. Because oranges float just beneath the surface and drift with the currents, the aerial photos capture a huge "map" marked out in orange-dotted lines. The lines trace water currents; interruptions in the lines show shipping lanes, deviations in the lines provide clues to submerged topographies, and so on. The shapeless murk of Biscayne Bay is illuminated and articulated simply by tracking the flow of waterborne objects through the system.

We live in a transparent, sea of cyberspace, and for the first time in history the flow of information through that sea is trackable and accessible over a global, growing Internet. This is a major development in the multi-billion year history of life and mind, and it is happening in our lifetimes. It is a big story, and a big scientific opportunity.

Second, as a sometime entrepreneur I think that tracking documents and information transactions will be a big business opportunity. When the information economy is as significant as the physical economy, "infonomic indicators" should be as important and as valuable to economists and market analysts as balance of trade statistics, the Dow Jones Industrial Index, etc. etc.

This raises further questions for rights-language developers. *Can ODRL allow content owners or superdistributors to claim ownership of valuable tracking data that are by-products of data-transactions, but not embodied in the rights-managed digital object itself?*

VIII. Tracking information flows: how?

Not surprisingly, my ideas about how to track information flows came from biology [2]. By exploiting the fact that each individuals genetic code is unique yet similar to that of close relatives, biologists have recently learned to reconstruct amazingly precise lineages of descent (pedigrees) going back hundreds of generations. These techniques have produced profound advances in biology, ecology, medicine, pharmaceutics, forensics, etc. Similarly significant advances would probably follow from a comparable system for reconstructing digital pedigrees of redistributed and evolving digital objects. After all, digital objects are increasingly the DNA of civilization.

One way to make digital objects trackable is record document transformations, reproductions, and the current context of use in a data field embedded within the object. Each time the object is accessed, we can check to see if the current context matches a previously stored fingerprint of the context, and if it does not, we can know that the object has been moved to a new context. In that case, we can append the new context fingerprint to the data field (thus preserving lineage information) and update our record of current context. In this way (and there are other ways) each digital object could have a family tree that would allow us to trace redistributed objects back, through all of its intermediary stages and users, to the original source. Then we can examine those data objects "in the field," or monitor their passage through mail servers, or have them periodically "phone home" to databases, and cross-reference with other data about purchases, purchasers, etc.

Does ODRL allow us to assert rights over, and prohibit tampering of, portions of a document that are intended to change, randomly or in a directed fashion, over time?

IX. Redistribution and the Privacy Problem

Document tracking also raises profound ethical issues. I suspect that a lot of redistribution tracking is already happening, but that it is unpublicized because document tracking invades the privacy of those who receive files as well as those who send them. It's a serious concern--suppose right-to-lifers used this methodology to identify and harass women to whom friends forward documents on abortion counseling? And recommended best practices are of little help: even if senders are informed about corporate privacy policies and allowed to specify the uses to which their personal data might be put, recipients of redistributed documents have no such choice or control.

I think digital rights management languages could be pressed into service here. Since the privacy problem has become a digital data problem, why not treat personal data as intellectual property owned by the people to whom it applies? If each us owned our personal data, each of us could use rights languages, copyright laws, and rights management systems to protect our privacy, fatten our wallets, and/or heighten public awareness of intellectual property law. The masses would benefit from the growing power of intellectual property law, and we could encourage people to make valuable data available and marketable. *(Does ODRL allow users to assert ownership of data generated by their handling of a given document?)*

A number of organizations have envisioned an anonymizing infomediary service, a "Personal Information Trust" (PIT), which collects, protects, and optionally sells anonymized personal information data in such a way that marketers could communicate with specific individuals (with consent under specified conditions) without learning the individual's identities, and in such a way that each individual could discontinue that communication at any time. Essentially the PIT would be a "go-between" or "Swiss bank account" that

could increase the value, and decrease the liabilities, of personal information by pooling information from diverse sources and by making a market for information buyers and sellers. [15]

The economics and ecology of the PIT would be quite interesting, because isolated snippets of data become more informative and therefore more valuable when they are commingled with other data in the PIT. This would allow the PIT to pay information deposits, and each information purchase would be add still more value to the PIT because data the information purchases are themselves valuable. A healthy PIT, like a healthy ecology, could actually "clean" the personal information environment by creating a value-gradient that would cause personal information to aggregate in the value-enhancing, privacy protecting, database, where it would earn money for the PIT and for the people it represents, through the sale of data and permission to contact targeted consumers.

While one can imagine many models for the governance, economics, and regulation of the PIT, the initial questions for rights language developers are clear. *Can ODRL be applied to data generated by information transactions between and among individuals, marketers? Does it allow individuals to specify the conditions under which they are willing to be contacted by marketers, or to let marketers or analysts make use their personal data (anonymized or otherwise) for other purposes?*

X. Evolving Digital Life

As I said earlier, living things don't just move from place to place. They reproduce, they spread, and they evolve. So far I have argued that digital objects can reproduce, can spread, and can be profitably tracked much like living things. I now want to suggest that its just a matter of time before they are "genetically engineered" to evolve and adapt through a process very much like natural selection.

Consider the case of a computer program that runs in "demo mode" for a certain number of minutes before demanding that the user purchase a "professional license". What is the right number of minutes? This might be hard to predict, and might vary from one market niche to another. But (1) if the number of minutes is controlled by a mutatable data field, and if (2) the number of minutes influences the probability that users will copy and redistribute the product, then the number of minutes should evolve, through random mutation and differential reproduction, toward values that maximize the likelihood of redistribution.

Thus, by putting functional aspects of a digital object under the control of mutatable code embedded in a frequently copied object, the conditions for natural selection could be created. We would want to select functional aspects that might affect the utility or attractiveness of a product, and we would want to constrain the degree of functional variation so that mutations could not have unacceptably negative (or fatal) effects. But even within such constraints there are many ways we might do this.

Of course, natural selection maximizes reproduction and this may not maximize purchasing, which is what product creators probably care about. But there are ways in which we might select mutations for purchase-encouragement rather than for copy-encouragement per se (see [2], columns 11 and 12).

The point is that in the long run, the difference between software and biology may become vanishingly small. Differential reproduction of inheritable characteristics – may eventually become another tool in the toolbox of the software engineer and the information marketer. If and when that happens, a new chapter in the billion-year history of life and life-like evolution may have begun. Indeed, in retrospect, we may conclude that the new chapter has already begun.

The last few decades brought us several digital revolutions, the open source software movement, the spam explosion, the copyright and patenting of DNA sequences, genetic algorithms, the onslaught of computer viruses and worms, and the emergence of a global information network. All of these things are driven by the "out of control" replication and propagation and evolution of digital objects, many with significant commercial value and social significance. It is the presumptive function of rights management languages to describe and facilitate the regulation or husbandry of these phenomena.

In this sense, rights management languages are themselves among the most interesting recent developments in the primal soup that constitutes today's information ecology. It will be interesting to see how well rights management languages can be designed for adaptive evolution.

References

[1] Schull, J. Method for encouraging purchase of executable and non-executable software. US Patent 5509070, 1992.

[2] Schull, J. Method for tracking software lineage. US Patent 6266654, 2001.

[3] Open Ebook Forum, Framework for an E-publishing Ecology, v0.78, 2000. www.openebook.org/doc_library/ecology/A%20Framework%20for%20the%20Epublishing%20Ecology.pdf

[4] Rozin, P., Schull, J. Adaptive-evolutionary perspectives and experimental psychology. In S.S. Stevens' Handbook of Experimental Psychology, R. Atkinson, R., Hernstein, G. Lindzey, & R.D. Luce (Eds.), New York: Wiley, 1988

[5] Schull, J. The View from the Adaptive Landscape. In Parallel Problem Solving from Nature, H.P. Schwefel and R. Manner, eds. Springer Verlag, 1991.

[6] Association of Shareware Professionals. History of Shareware. http://www.asp-shareware.org/users/history-of-shareware.asp 2000.

[7] Dawkins, R. The Selfish Gene. Oxford: Oxford University Press, 1976. Chapter 11.

[8] Kelly, K. Out of Control: The New Biology of Machines, Social Systems, and the Economic World, 1994. Reading, Mass.: Addison-Wesley, 1994. http://www.kk.org/outofcontrol/contents.php

[9] R. Mori, and M. Kawahara, Superdistribution: The Concept and the Architecture. THE TRANSACTIONS OF THE IEICE; VOL.E 73, NO.7 JULY l990 Special Issue on Cryptography and Information Security. http://www.virtualschool.edu/mon/ElectronicProperty/MoriSuperdist.html#Mori

[10] Cox, B. Superdistribution. *Wired*, 2.09, Sept 1994. http://www.wired.com/wired/archive/2.09/superdis.html

[11] Cox, B. *Superdistribution*. Reading, Mass.: Addison-Wesley, 1995

[12] Weber, R. The Superdistribution Chronicles, Pt. 1. http://robertweber.typepad.com/rightsmanagement/2005/02/the_superdistri.html

[13] Doctorow, C. *Down and Out in the Magic Kingdom.* New York: Tor, 2003. http://www.craphound.com/down/download.php

[14] King, S. Riding the Bullet.
http://www.simonsays.com/content/content.cfm?sid=33&pid=479688
[15] OpenPrivacy.org, An Annotated OpenPrivacy Bibliography
http://www.openprivacy.org/bibliography.shtml

Jonathan Schull received a Ph.D. in Biological Psychology from the University of Pennsylvania in 1980, taught at Haverford College until 1992, and then gave up tenure to found SoftLock Services, an early rights management company, that later changed its name to DigitalGoods.com. In 1999, Schull drafted the Open Ebook Forum's "Framework for the Epublishing Ecology. By 2000, when Stephen King published his e-book "Riding the Bullet", DigitalGoods was a 75 person company listed on the Nasdaq stock exchange. In 2001, the company closed its doors and liquidated its assets, including patents Schull authored in 1992. Schull now teaches Human Computer Interaction in the Interactive Media group, in the Information Technology Department of the Rochester Institute of Technology, and continues to pursue his long-standing interest in the dynamics of intelligent and adaptive networks.