Koepsell, David. "Introduction." *Innovation and Nanotechnology: Converging Technologies and the End of Intellectual Property*. London: Bloomsbury Academic, 2011. 1–23. *Bloomsbury Collections*. Web. 26 Apr. 2023. <<http://dx.doi.org/10.5040/9781849664783.0006>>.

Downloaded from Bloomsbury Collections, www.bloomsburycollections.com, 26 April 2023, 12:45 UTC.

Copyright © David Koepsell 2011 2011. You may share this work for non-commercial purposes only, provided you give attribution to the copyright holder and the publisher, and provide a link to the Creative Commons licence.

INTRODUCTION

Nanowares: Science Fiction Futures and Present Potentials

Consider this: everything that surrounds you is matter. Matter comprises every object in your sight, everything your body is composed of, processes, excretes, and thinks. Yes, every thought in your head is dependent upon matter as well, and as we know because of Albert Einstein, even energy is matter in another form. The atomic theory, which posited that the universe is composed of very tiny bits and pieces, each of which is further composed of tinier bits and pieces, is one of the most successful theories in science, having been confirmed by a hundred years of observation. The bits and pieces that compose the universe are the atoms, of various sizes and qualities, that we are now all familiar with from the Periodic Table of Elements. In the mid-twentieth century, as the atomic theory continued to be successfully confirmed, and humans grew increasingly capable of seeing and doing things with the world of the very, very small, some forward-thinking scientists began to wonder whether there were physical limits to the ability to manipulate atoms, and if not, what such limits might mean for our abilities to remake the world around us.

If, they thought, we could manipulate matter at the atomic level (at the ‘nanoscale’, a scale measuring one to one hundred nanometers, or one-billionth of a meter), then conceivably, we could develop very tiny machines that could in turn build things for us at both the micro- and macro-scales. Imagine if you could construct nanoscale robots (essentially, the size of large molecules) and instruct them to build a computer, or a car, or anything you desired, from readily available raw materials. Or consider the possibilities of using such nanoscale machines to combat diseases or parasites, viruses or cancers in the body, eliminating the need for the scattershot or blunderbuss approaches of much of modern medicine. Or perhaps we could develop radical new materials, incredibly strong, capable of mending themselves, light, efficient, or even imbued with properties we only dream of in science fiction. The possibilities of nanotechnology, which is the application to engineering artifacts of our increasing knowledge about physics at the nanoscale, are wide open.

The history of technology in general is a history fueled largely by the science of slow miniaturization and incrementally better tools for manipulating

things at ever smaller scales. The past hundred years of electronics demonstrate this evolution, as vacuum tubes became transistors, and as silicon chips became quantum processors. The science of the small, from which the technology of computing has remarkably benefited, is now reaching out beyond the two-dimensional world of silicon wafers, and promises to remake the world of more-ordinary things, the hardware that composes the rest of our technology and human artifacts, and that may revolutionize the way we interact with our environment and tools. The history of computing serves as not only a catalyst for the coming age of nanotechnology, in which we will be able to begin to manipulate the world around us so that we can build remarkable new tools literally from their constituent atoms, but also as a suitable departure point for a discussion of the ways in which nanotech will alter the ways we innovate, and how we both encourage and protect innovation.

For about the past two hundred years, we have split the world into two. On the one side stood our artistic and creative artifacts, and on the other our utilitarian inventions. This dichotomy has been increasingly important, and lately challenged, as our artifacts became more closely tied to digital expressions and electronic media. Computerization, brought about largely through substantial breakthroughs in miniaturization, ushered in a new age of innovation in which machines and aesthetic expressions began to merge. The legal paradigm encompassed by intellectual property (IP) law, and embraced by a larger culture, began to break down.¹ Machines and aesthetic expressions no longer seemed so distinct. How can we learn from the mis-steps and impediments posed by the failing two-world paradigm, and build new institutions, both legal and cultural, that are not only philosophically sound, but that can encourage the sort of innovation promised by futuristic nanotechnology?

This book comprises an argument that (a) current schemes of IP protection are not only pragmatically incapable of being applied to nanotechnology,² but also theoretically inadequate to promote modern innovation, and (b) nanotechnology reveals that our relationships to our artifacts have been misunderstood, and poorly applied through laws that were once considered necessary to promote innovation, but which now prove to be only impediments. This argument is made by looking first at general trends in the history of innovation, and the technologies and institutions that have enabled and sped it in the past hundred years,

converging now with nanotechnology in its broadest sense; and then looking at the first principles involved in our theories of IP, and criticizing their bases as well as effects. Finally, a new paradigm will be offered, based upon methods and processes being adapted in both the technologies of innovation and new forms of protection being developed at the grassroots, eschewing old forms of monopoly, and embracing both the spirit and methods of open science and innovation.

But first, we'll look at the general trends of industrialization, miniaturization, delocalized production, and of course, the 'convergence' that all of this is centering upon in modern and 'futuristic' nanotechnology.

From Arrowhead to Atom

Often, when we hear the term 'artifact' we think of ancient tools discovered in strata of ancient soil. But the term artifact simply means anything purposely created by humans in some enduring medium. Thus, a song is not an artifact when sung, since it drifts away into the aether never to be heard beyond its immediate performance, even though it is intentionally created by humans. However, an ancient recording of a song on a 78 rpm platter would be an artifact. A skeleton is not an artifact, until someone carves something decorative or useful upon it or arranges it into some purposeful position.³

Proto-human artifacts date back millions of years. The archeological record holds increasingly older surprises as we find that ancient humans have been turning found objects into tools at earlier dates than we even recently suspected. The first human artifacts came about the moment some early proto-human mixed labor with a found object with the intent to create something new. This is the genesis of craft, of art, of technology. Intention, as we shall see is critical, and is what makes artifacts different from accidents.

The history of human art slopes gradually upward from the first-fashioned arrowhead or similar tool, slowly tracing a halting curve toward the modern industrial age, becoming hyperbolic in the last fifty or so years. Most of the milestones upon that curve involved changes in the manner and means of production. The course of that history will be sketched here only in brief, and in broad strokes, because the way we evolved from chipping away at flint and bones, to being able to assemble machines from the atomic scale on up, illustrates the critical change that is posed by

futuristic nanotechnology to our assumptions regarding our relations to our artifacts, and the ways we have come to encourage innovation through institutions.

Ancient technology was decentralized. Even while the manner of producing woven baskets, clay pots, ceramics, and eventually metal-ware demanded increasing skills and specialized knowledge, the artisans who made the crafts that formed our most ancient technologies created their goods by laborious, local, and more-or-less *ad hoc* manners. In other words, when an arrowhead was needed, it was fashioned most likely by the end-user, and perhaps in some surplus as necessary for an upcoming hunt. Perhaps surpluses were traded for goods produced by those with other skills and arts, and some form of ancient trade began to create specialties. But this sort of barter cannot be mistaken for a market, and technologies not otherwise impelled by more than mere necessity (no profit motive, for instance) grew only incrementally with the advent of new necessities. Thus, as ice ages waxed, new needs for clothing impelled the creation of new, local arts that could meet new needs and sustain life.⁴

This mode of production and dissemination of technology marks the first major epoch of technology. Driven by necessity, requiring only modest skills, and only very little in the way of specialization, the first human artifacts changed very little over the course of humanity's first million or so years. Up until about 10,000 years ago, in the Paleolithic or 'Stone' age, human tools were created as needed, from simple and available materials, and show only a modest degree of slowly increasing artistry. Human arts in general seem to have fulfilled the distinct purpose of survival for that first million years, aiding with the hunt, both materially and spiritually, as humans wandered from place to place. It was only with the invention of agriculture, and the advent of the Copper and Bronze Ages about 10,000 years ago, that human technology began to advance significantly, and the rate of change began to increase.

With agriculture came an end, more or less, to nomadic life. And with a more stable lifestyle, and the beginning of what we might call *culture* or even *civilization*, came the ability to specialize tasks more discretely, and to build more permanent means for creating artifacts or techniques. It is interesting to note that there are two very different and ancient means of creating an artifact, both of which emerge very early on, but one of which comes to surpass the other. Things can be made by either (a) taking something away from an

existing thing (by chipping or carving bits away), or (b) ‘building’ a thing from some formless mass. For example, a piece of flint can be chipped away until an arrowhead remains, or a clay pot can be molded from formless clay into a thing with form and function. Most of the artifacts we are surrounded by today fall into the second category, and it is the paradigm behind many ideas for creating nanowares, by which anything could be ‘built’ from the bottom-up, and eventually atom-by-atom. These two basic means of producing artifacts will be discussed in more detail as we begin later on to delve a bit into the metaphysics of nanotechnology, but it’s also worthy of note that the move from *a* to *b* was impelled and made possible by a shift in the necessities driving the creation of artifacts, and the modes and means of production that became available as humans began to settle into communities.

As nomads settled into stable groups, villages, towns, cities, etc., with cultures, so too did technique and artifacts begin to flourish. Without the challenge of everyday necessity impelling and permitting only that technology which affords immediate survival, artisans and craftspersons could commit more time, more resources, and more thought to developing tools meant to fulfill not just needs, but *interests*. Moreover, while the flexibility of crafts that rely upon found objects becoming fashioned into useful goods is low, new materials, primarily metals, pottery, and ceramics make building things from scratch a preferred and much more flexible mode of creation.

The techniques available for crafting new things out of flexible media (as we’ll call, for now, clay, metal, and other media from which things can be built ‘up’, rather than cut ‘down’) could further be perfected as leisure time increased. As artisans became more adept at a particular, chosen technique, specialization became possible. Categories of craftspersons could thus improve in their practiced techniques as artisans need not be ‘jacks of all trades’. All of which forms the basis for the development of an economy, as specialists learn to trade and barter their arts with others, producing their creations in strategic surpluses, affording themselves the option of trading their goods for those of others. Around 5,000 years ago, the first writing systems emerged out of the need to track trades of goods, and the wheel became adapted for carrying goods across distances, between various trade centers, expanding the scope of what we’d now call ‘markets’ for goods.⁵ The emergence of trade and barter, the development of dispersed markets,

and the luxury of specialization required certain manners of production, storage, and dissemination of goods. Surpluses became a strategic necessity, replacing the production of things as needed with a system by which one could begin to profit, and accumulate wealth, enabling the slow emergence of more modern economies.⁶

With specialization, crafts could develop into trades. Trades then developed their own cultures, policing their ranks, and ensuring market domination over a territory. This marks the next stage of the development of human technologies and markets, as arts became specialized, and the first, nascent form of ‘IP’ protection began to emerge. Secret-keeping, enforced through trade associations and guilds, is the first means by which those who practiced useful arts sought to control their markets. As technologies became complex enough to require training to produce or practice, and too difficult or time consuming to easily ‘reverse-engineer’, secret-keeping became an effective means to ensure that trades could dictate higher prices for their goods by policing the market and preventing undue competition. This in turn propelled the continued perfection of technologies, to prevent reverse-engineering, and to help ensure market domination by a particular trade.⁷ Specialization and trades came to dominate the emerging market for artifacts for much of the past two to four millennia. But the growth of a new manner of acquiring knowledge, beyond the scholastic method (in which masters taught apprentices, and knowledge was kept obscure and secret from the uneducated masses), began to succeed where scholasticism failed, and thus began to assume primacy as humankind’s foremost method of inquiry. Science slowly emerged as the dominant epistemological paradigm in the West during the Renaissance and Enlightenment, and led Europe and the Americas into the industrial revolution beginning in the late 1700s.

Science demands unfettered inquiry into the workings of nature, and replaces the confidence previously demanded over rote knowledge with a practiced skepticism, and ongoing investigation. With the rise of the age of science came the need to develop new means of treating information. Scientific investigations conducted by ‘natural philosophers’ could only be conducted in full view, out in the open, with results published in meetings of scientific societies and their journals. Supplanting secret-keeping and obscurantism, the full sunlight of public and peer

scrutiny could begin to continually cleanse false assumptions and beliefs, and help to perfect theories about the workings of the world.⁸ Science demanded disclosure, where trades and arts often encouraged secrets. And so as natural philosophers began to disseminate the results of their investigations into nature, new forms of trade, art, and industry began to emerge, as well as the demand for new means of protection in the absence of secrecy. Thus, as the scientific age was dawning, and helping to fuel a new technological revolution, *modern* forms of IP protection such as patents and copyrights emerged as states sought to encourage the development of the aesthetic and useful arts. By granting to authors and inventors a monopoly over the practice of their art, as long as they brought forth new and useful inventions (or for artistic works, as long as they were new), nation states helped to attract productive and inventive artisans and trades into their borders. These forms of state monopoly also enabled further centralization of trades and industries, as technologies now could become immune from the possibility of ‘reverse-engineering’ and competitors could be kept at bay by the force of law. This sort of state-sanctioned centralization and monopoly helped build the industrial revolution (by the account of many historians and economists, although this assumption has lately been challenged) as investors now could commodify new technologies free from the threat of direct competition, secure in the safe harbor of a state-supported monopoly over the practice of a useful art for a period of time.

In many ways, traditional IP was (and is) deemed vital to the development of large industries and their infrastructures, and to the centralized, assembly-line factory mode of production that dominated the twentieth century. With the benefit of a state-sanctioned monopoly, industry could build sufficient infrastructure to dominate a market with a new technology for the duration of a patent. This confidence assured investors that there would be some period of return on the investment in which other potential competitors are held at bay, at least from practicing the art as claimed in the patent. Factories could be built, supply chains developed, and a market captured and profited from, and prices will not be subject to the ruthless dictates of supply and demand. Rather, because of the luxury of a protected market during the period of protection, innovators can inflate prices to not only recoup the costs of investment, but also profit as handsomely as the captive market will allow.

For most of the twentieth century, IP allowed the concentration of industrial production into the familiar factory, assembly-line model. Even while the knowledge behind new innovation moved eventually into the public domain as patents lapsed, during the course of the term of patent protection, strictly monopolized manufacturing processes and their products could be heavily capitalized, and substantial profits realized, before a technique or technology lost its protection. But the modes and methods of manufacturing are now changing, and the necessity of infrastructural investment is also being altered by the emergence of new means of production, including what we'll call 'micromanufacturing', which is a transitional technology on the way to true MNT (molecular nanotechnology), and is included in our discussions of 'nanowares'. Essentially, assembly-lines and supply chains that supported the huge monopolistic market dominance models of the industrial revolution, well into the twentieth century, are becoming obsolete. If innovation and production can be linked together with modern and futuristic breakthroughs in micromanufacturing (in which small components can be fabricated and produced *en mass*, cheaply) and eventually molecular manufacturing (in which items are built on the spot, from the ground up, molecule by molecule), then we should consider whether the IP regimes that helped fuel the industrial revolution are still necessary, or even whether they were ever necessary at all. Do they promote new forms of innovation and production, or might they instead stifle potentially revolutionary changes in our manners of creation and distribution?

What we see when we look at the broad strokes of the history of artifacts, their innovation, production, and distribution, reveals clear trends. There is a general move from local, unspecialized, and decentralized to centralized and specialized, facilitated by institutions and states. There is a general move from use of available objects that can be fashioned into useful tools to developing tools from the 'ground up' from flexible media. There is a general move from the practice of arts in a scholastic mode to the incorporation of the methods of sciences into the process of innovation. There is also a general trend of miniaturization, facilitated most recently by the silicon revolution and computerization, which, as we'll see in the next section, parallels and predicts much of the coming issues this book is concerned with in the coming nanowares revolution, including prominently the role of IP regimes, and their effect on innovation.

A Very, Very Short History of the Very, Very Small

Precipitated first by the development of electricity, and the gradual move from a steam-powered world to an electrical one, miniaturization of components enabled both energy savings and smaller, more portable goods. While computing might have taken a mechanical turn, as the first calculators and rudimentary computers were mechanical (like Charles Babbage's Difference Engine), the success of new means for distributing electrical power achieved with Nikola Tesla's alternating current created a market demand for electrical goods, which in turn led to the development of vacuum tubes, then transistors, and ultimately silicon chips. Each successively smaller generation of products could be delivered to more customers, at more affordable prices, with less energy demand per unit, making operating costs also more affordable.⁹

Miniaturization began, of course, before electricity, primarily with timepieces used for navigation. John Harrison's famous H-series chronometers were built to meet a challenge posed by the British Crown, that of solving the biggest and most dangerous problem of naval navigation: longitude. Determining latitude was relatively simple given the angle can be determined from the pole by measuring the angle of relative declination of the pole star. But determining longitude was a singularly troublesome matter without accurate timepieces. While relatively accurate timepieces existed by the early 1700s, none were portable for the purposes of navigation at sea, as the shifting caused by the waves disturbed the mechanisms sufficiently to cause the best timepieces that existed to gain or lose time too greatly to be effective for long trips. Harrison solved this problem by developing very small, precise, spring-driven mechanisms for his timepieces. Miniaturization of chronographs, and their development into necessary fashion accessories, pushed the limits of mechanical miniaturization to their limits.¹⁰

In fact, clockwork mechanical computers were used for complex problems well into the Second World War. The larger these machines grew, the more power necessary to operate them, and thus steam, and electricity-powered devices were envisioned and built that could do rapid, complex calculations. To provide for storage and retrieval of information so that a general-purpose computer could run various different programs, punch cards and other mechanical means of read-only memory were incorporated into the earliest computers. The Jacquard loom is essentially such a computer, capable of

reading cards with a weaving sequence programmed on them, and weaving the sequence automatically as the loom is powered. The electrically powered Mallock machine from the 1930s is an example of a mechanical, analog computer.¹¹

The power and precision necessary for complex mechanical computers (especially analog ones – involving ranges of values *between* 0 and 1) helped push the development of digital computing and the adoption of logic gates in the form of vacuum tubes. Whereas reprogramming a mechanical computer could be accomplished through the introduction either of new mechanical elements or through punch cards, a truly flexible, random access machine, or ‘universal computing machine’ of the type theorized by Alan Turing, was made possible with digital, electronic computing. The machinery of a digital, electronic computer is its memory, in which both the data and software for manipulating that data can be updated without reconstructing the machine itself. The speed, flexibility, and eventually the size constraints posed by mechanical computation could be overcome with the introduction first of vacuum tube logic gates, storing data as ‘1s and 0s’ (really, differing voltage states representing the numerals 1 and 0), and then as components became smaller and smaller, in transistors, and finally, silicon chips.¹²

Driving the need for ever smaller components were two factors: energy consumption and portability. The first vacuum-tube digital computers were monstrous, room-filling beasts that pumped out so much heat that they required enormous cooling mechanisms, and the combination of energy used by the computer and the air-conditioning made them prohibitively expensive to operate. Only wealthy universities, government facilities, and corporations could afford to own a computer and to operate one regularly. These sorts of computers were sufficient for the mostly military needs they fulfilled during the Second World War, but after the war, scientific and corporate demand for better, cheaper, and faster computing created a market for smaller, more efficient machines.¹³

The move to transistors and away from vacuum tubes as the means of computerized logic allowed for smaller, cooler, and more portable machines. The miniaturization of electronics helped fuel the consumer electronics revolution that drove an increasingly large amount of the western (and Japanese) technological economy through the 1950s, 1960s, and early 1970s. Transistors still relied on conductors that conducted heat poorly, and

that impeded current, resulting in energy consumption and unreliability. Only with the development of the integrated circuit, and logic boards whose primary medium was silicon, could today's fast, cool, and relatively cheap computers be designed.

Energy efficiency and portability drove miniaturization in computing, and within a hundred years, the wildest dreams of Charles Babbage and others who were designing steam-driven, mechanical computers were exceeded by ubiquitous devices like modern cell phones, whose computing power is hundreds of times more than the best personal computers just twenty years ago. Silicon proved to be a magical medium, allowing the design and production, through photographic/chemical etching, of intricate and ridiculously small binary churning integrated circuits. The chips that are designed today have components so small you may need an electron microscope to see their full workings.¹⁴

Silicon has allowed for computerization to bloom, meaning that the range of human creativity that can be carried out by inputting data, manipulating logic gates, and outputting information has unleashed a torrent of new types of creativity and innovation. Take, for example, written expressions. At one time, shortly after the initial invention of the printing press, the ability to spread written communications was greatly increased, even while the producers of printed texts maintained a fairly secure monopoly over their trade through secrecy (printing presses were initially difficult and expensive to reverse-engineer, much less mass produce), and through intimidation as printers guilds developed. Printing technologies remained expensive well into the twentieth century, even when the technique behind the technology was well-known. The costs associated with acquiring a press, and the time and energy that were required to create original press runs of a book, newspaper, or pamphlets, kept the technology in the hands of those who could afford a significant capital investment. Now, because of the ubiquity of computers, and their ever-decreasing costs, creating printed goods is not much more expensive than the costs of the paper and binding. Anyone can, in effect, become a publisher and even, thanks to several small startups specializing in 'printing on demand', deliver bound, professionally printed books to anyone willing to pay the prices of shipping all over the world. If one forgoes the printed text, then publishers can instantly distribute their written expressions almost without cost anywhere that has internet access.

Or consider movies. As with printing, the tools of production were once prohibitively expensive for all but those with healthy amounts of capital to produce and distribute films. The cameras, film stock, and developing the negatives were all quite expensive, at least for professional quality productions worthy of wide distribution, and the infrastructure of film production and distribution were also controlled by trades with significant organizational inertia and power. But again, silicon and modern digital computers have liberated the means of both production and distribution, empowering the movement from idea to expression to become both rather simple and cheap, available to anyone with the creative wherewithal and just a little up front money. The primary requirement for realizing one's creativity is now *intellectual capital*. The same story applies to music, as well as to software itself, the engine behind the computer revolution.

These stories illustrate the trends that are broadly painted above: the move from art to trade, from delocalized to centralized, from cheap to expensive, and then back again. All of the modern computer and internet revolutions are built upon the move toward smaller, cheaper, and more energy-efficient modes of computing. Advances that have created modern computing are now being turned toward manipulating more than just silicon. These general trends have come, in a sense, full circle with *aesthetic* expressions of a certain type: those that can be realized digitally, with silicon-based computers and networks. The same general trends will mark our transition from the current mode of creating *objects* to the future or nanowares. When objects become as easily programmable and cheap to make as silicon-based expressions are, then anyone with an idea will be able to realize it in matter as easily as today's innovators in home-made texts, films, and songs. We should take heed then of the role of IP regimes in the computer revolution, and consider what we might learn from IP's role in the computer revolution to ease our transition to the nano-future.

Intellectual Property: A Concise History in Computing Technologies

Software has posed a problem for legal scholars and jurists since its inception, at least in the framework of IP that we have lived with for the last several hundred years. When IP began to make its debut a couple hundred years ago with statutes or acts of sovereigns in Europe giving publishers

or inventors monopolies over their works, the distinction between a book and a printing press was clear. Books were passive machines for storing information, whereas a printing press was an active machine for making books. No one could have conceived that one day these two seemingly very different types of artifacts would converge into one powerful new technology. A computer fits neatly into our preconceptions of what a machine must look like, especially when computers still consisted of glowing tubes, blinking lights, punch cards, and other clearly machine-like appurtenances. Software, however, is another story. Software was not originally very different seeming than other parts of a computer. In fact, consider a computer that is programmed only for one task, like multiplication. In such a machine, the software is hard-coded into its logic gates, and those logic gates are clearly parts of a machine. The software is simply the hard-coded elements of a larger machine. Thought of this way, the software of any machine (even the design of a particular steam engine, for instance) is the arrangement of its specific parts in such a way that makes that machine perform its function as it was programmed (designed) to do. But the design of a machine is not software as it has *now evolved* to be. Re-programmable machines can perform new functions with the introduction of new instructions, and machines like these have existed before modern computers, pushing our definition of what constitutes a computer back further than we'd at first expect. We could consider inventions like the Jacquard loom to be early computers. These looms helped automate and standardize the process of weaving specific patterns into materials. When done by hand, weaving requires careful attention to when to throw the shuttle, and when to raise or lower the harness, producing a specified pattern. Human operators would read written notations which indicated when to do which action so that a pattern could be produced, and reproduced. Think of this written notation like sheet music. In 1801, Joseph Jacquard automated this process by fitting a loom to 'read' punch cards with holes that caused the raising or lowering to occur automatically, rather than requiring a human intermediary to read the notation. The machine reads it. The human-readable notation is substituted with a machine-readable software that eliminates a human step from the machine, speeding up the process, and helping to eliminate errors.¹⁵

The same thing happened a hundred years later with 'player' pianos, which anyone who has seen a Hollywood Western knows operates by 'reading'

a roll of paper with punched-holes, eliminating the human intermediary called a piano player, and allowing the machine to read the music in software form. The technological evolution from Jacquard looms to player pianos is obvious, as is the move from hard-wired computers to punch card reading computers, but who knew that a legal case involving player pianos would make copyright unavailable to software until the 1970s as well as making it unavailable to vinyl records and audio tapes?

Expressions versus Machines

The rule used to be simple: in order to be an ‘expression’, and thus amenable to copyright protection, a writing had to be directly perceivable by a human. Words on a page, paintings, photographs, statues, and even sheet music would qualify under the ‘direct perception’ test. But player piano rolls were another story. In the *White-Smith Music Publishing Co. v. Apollo Co.* case of 1908, the Supreme Court of the United States made explicit the ‘direct perception’ rule, and thus excluded from copyright protection any expression that could not be directly perceived (or rather, understood, because humans can of course perceive holes).¹⁶ As a result of this ruling, phonograph records, audio tapes, and other modes of recording information in ways that were not directly *understandable* by ordinary humans could not be protected by copyright. It took an amendment to the Copyright Act in 1972 to change this, and the direct perception test was finally abandoned.

The 1972 Copyright Act corrected a logical error in the law, which survives still in the dichotomy between patent and copyright, that distinguished between types of expression based upon human faculties. This error seems to have been based upon the notion that if humans cannot immediately understand an expression, it isn’t really an expression. The illogic of this point of view is now quite clear. Expressions in foreign languages cannot be immediately understood by those who don’t speak those languages. Some humans could likewise plausibly train themselves to understand mediated expressions like player piano rolls, just as they can learn to read sheet music. But sheet music, writing, the grooves on a vinyl LP, the patterns of 1s and 0s on computer tapes and discs are all media for transmitting ideas, and are all likewise expressions. The *purposes* of different types of expressions may differ, but if copyright law is meant to protect human expressions, then excluding a range of mediated expressions just because they generally

require some further translation by another machine for their enjoyment makes no logical sense. The 1972 amended Copyright Act was doubtless impelled by pressure from music publishers and record producers to give copyright protection to their recordings, and to help ensure future profits from a growing base of music consumers, but the amendment had the likely unintended consequence of fixing the metaphysical absurdity of excluding machine-readable expressions from the broader set of *expressions* in fixed media.¹⁷

What is an expression anyway? How does an expression differ from a machine, and why does the law create the dichotomy between the subjects of patent and copyright? Before the US Congress fixed the artificial distinction among machine-readable and human-readable expressions in 1972, computer software was deemed patentable. The instructions in a computer, which open certain logical gates at certain times and close them at others, causing *computation* and outputting results, are *parts* of a machine. A hard-wired computer, as discussed above, can only do one function. But the same architecture could be re-wired to perform different functions, and modern computers are useful precisely because the instructions for opening and closing logical gates can be changed by the introduction of new software. The computer then becomes, technically, a new machine when the new software is input. It is one machine when it is off (one capable of performing any number of functions), another machine when it is turned on and the underlying software is activated (like an operating system), and another machine as new pieces of additional software are ‘loaded’. In each case, the software that directs its actions is part of the total machine, making the computer an enormously (if not infinitely) reconfigurable machine, capable of taking on new forms. Which part of the machine is the expression? Should software be properly considered an expression if, as we can see, it is part of a machine?

Yes, software is an expression. The error of the law, corrected only partly still, is in distinguishing between a certain type of expression and a machine in the first place. At least with the amended Copyright Act, the artificial distinction between those expressions that are machine-readable or mediated versus those that are ‘directly perceived’ by humans has been knocked down. But in what sense is it proper to exclude machines from the category ‘expressions’ at all? A machine fails only to be a certain *type* of expression, just as punch cards failed to be a certain *type* of expression.

After all, a punch card in use is part of a machine, as is other software. But I have argued at length and for some time that the distinction among machines, books, songs, etc., is artificial, and not founded upon any sound metaphysics. In *The Ontology of Cyberspace*, I first argued that anything that is man-made and intentionally produced is an expression. Thus, the whole range of human artifacts, and even some less-than-clearly artifactual things (like a spoken word) are properly called expressions. The uses of these various types of expressions clearly differ. Some are meant to give us aesthetic pleasure, other expressions serve more ‘utilitarian’ ends. But the largest distinction underlying all of IP law, which says that machines are different than songs, is ontologically mistaken. It took the flexibility of computerized media to reveal this faulty ontology, and it still has not been corrected in the law, despite the many practical problems this has caused in the software industry.

The IP Mess in Software

The development of IP, once encouraged as a useful means of promoting innovation, creativity, and disclosure, may be doing the opposite in software. The first IP protection afforded software was patenting. But patenting is expensive to pursue, costing years of precious time before completion and often thousands of dollars. Copyright is free, you can do it yourself. Every author instantly owns the copyright of their own original works, even without affixing the © mark with their name and year of production to their work, if the work is ‘fixed’ in some tangible medium. Once the amended US Copyright act specifically included what ought to have never been excluded – expressions that were machine-readable – software authors began to copyright their works rather than patenting them. But this posed both a practical problem and a metaphysical conundrum, all of which pointed even more clearly to an underlying fault with IP law.

The realms of patents and copyrights were once held to be mutually exclusive categories. Something could never be both patentable and copyrightable. A wealth of cases and the well-established practices of patent offices stood for the proposition that a utilitarian object, as long as it was new, non-obvious, and useful, could *not* receive a copyright. Moreover, an aesthetically oriented ‘work of authorship’ could not be patented. Never mind that aesthetic pleasure is a use. I have often wondered whether the distinction is rooted in Calvinism, as the so-called ‘useful arts’ are rewarded

with strong (though relatively brief) monopolies, and governmental intervention, whereas authors and artists are more or less left to their own devices for the protection and enforcement of their creative turf. But this strange distinction among legitimate uses resulted in a true conundrum with software, leading to a breakdown in previously mutually exclusive categories, and causing some to wonder whether software was some sort of new object that did not fit neatly into the previously clear categories of patent and copyright. I have argued that, in fact, these categories were poorly construed from the start, and that all expressions ought to be treated as a kind, and each new man-made object falls somewhere on a spectrum between ‘mechanical’ usefulness and other uses. Software has revealed for us the faulty ontology underlying all of IP law, and continues to wear away at its foundations.

Yet today, software patents are multiplying, seen as a favored apparatus especially for wealthier software companies to protect their innovation through strong monopolies, and to expand patent portfolios with which potential competitors can be excluded from the use of certain algorithms. Much of the value of software giants like Microsoft and Apple rests upon the size of their portfolio of patented algorithms. With these, and a cache of patent attorneys at their disposal, these companies can keep smaller competitors at bay. Patents are ideal for this given that there is no exception for ‘honest’ infringement. In other words, if you happen to develop a neat and potentially profitable algorithm, and never knew it was part of some previously patented code, you’re out of luck. You cannot use it. You must either ‘work around’ it, or license it from the original patent holder in order to incorporate it into your invention. This is in stark contrast to copyright law which does allow for the independent, innocent creation of a substantially similar expression without infringement. This is perhaps why software patents have grown in popularity with those who can afford them, and why many smaller, newer companies are opposing software patents and working outside the system. Having seen the mess that software patents can cause, and the roadblocks to innovation that they feel are posed by patents on algorithms, many software developers are choosing either to rely once again on the ancient mode of protection known as trade secrets (Google’s search engine is built on trade secrets) or using new forms of open development, including Open Source licensing, by which profits are gained not through an artificial, governmentally granted monopoly over an art,

but rather through quick innovation, customer service, and moving good products swiftly to market.

Of course software is patentable subject matter, but then, if we are to be fair, so is James Joyce's *Ulysses*. When it was produced, it was new, non-obvious, and useful. It was an object, created by man, and a composition of matter, or even arguably a process. Under the current, broad interpretation of patentable subject matter, anything on earth made by man ought to be susceptible to patent. While some would argue that *Ulysses* lacks 'usefulness', this is an impossible position to defend for a work that has been found useful to so many consumers willing to pay to own it, to submit to the difficult task of reading it, and sometimes even enjoying it. The use is a primarily aesthetic one, but then so too are the uses of many machines at least partly aesthetic. And so while software patents are debated in the markets, in the courts, and in the media, none of these debates seem to get to the heart of the debate: why do we distinguish between types of expressions, when software has revealed for us the flaw in the distinction between patent and copyright? Why not have a single, unified scheme of IP protection that recognizes that all man-made objects intentionally produced (an important *caveat*, because some man-made objects are accidents) are expressions, and distinguishing amongst types of expressions is no longer meaningful?¹⁸

This has troubled me for the past decade, and the world seems no closer to resolving this problem, entrenched as we are in traditional forms of IP protection despite their growing obsolescence in the face of new technologies. If software didn't destroy the old IP paradigms, it has surely undermined them. Nanotechnology is poised to finally do what software failed to do: destroy IP for good.

Nanotechnology, Microfabrication, and the End of Scarcity

The trends broadly painted above illustrate a number of convergences, some of which will draw us into societal conflicts that may either propel innovation in a new direction or stymie growth and limit potentials. While much angst has been expressed about the safety and security risks posed by nanotechnology, these are more or less science fiction scenarios that do not concern everyday researchers in the field who are battling with the enormous physical and technical hurdles of developing anything like the

futuristic nanotechnology in Michael Crichton's *Prey* and similar scenarios. Self-assembling nanobots are decades away, but the possibilities posed by the trends of delocalized fabrication, nanotech's precursor, could well be impeded by traditional IP practices and norms. The liberating possibilities posed by the ability to make nearly anything, anywhere, whether through current technologies (like Fab Labs or cheap, open source 3D printers, discussed more in subsequent chapters), will not be realized under current IP regimes, for many of the same reasons that software patents are impeding innovation in that marketplace. The challenge initially posed by the anti-innovative effects of software patents will leak into the world of 'physical' goods as all objects around us become more like software. This could be a tragedy, given the vast promise that new technological trends hold for solving all problems related to scarcity.

Scarcity is at the heart of the problem. The world of physical goods has long enjoyed robust markets, realizing profits for craftsmen and other producers, and abundant resources for consumers where those products were never afforded any IP protection. Grain, fruits, vegetables, the simplest farm implements, tables, chairs, knives, spoons, and nearly every other 'necessity' of life no longer enjoy any IP protection, and yet the markets for these types of goods survive. Supermarkets teem with the exchange of unprotected goods, whose value stems from traditional, well-understood mechanisms of supply and demand. Artificial scarcity in a truly competitive marketplace means that a competitor can undermine another's prices and defeat dishonest attempts to manipulate supply. Thus, smart farmers produce as much as the market will bear, and not much more. Warehousing perishable products won't work, and in a free market (though none really exist in agriculture as states create mechanisms to support local producers, thus defeating the forces of supply and demand to some extent) overabundance is no better than scarcity, as the price will surely plummet, and everyone will suffer. But markets for other types of goods are constantly manipulated, even as are those for agricultural products, not just by market players, but by governments seeking to provide advantages to their own producers, even where doing so may ultimately harm consumers.

IP is a form of artificial scarcity, created by governments to provide local producers with a market advantage that skews free markets. There is a very good reason proposed (but not proven) for this, at least from the viewpoint of encouraging innovation. If you don't think people will

innovate unless scarcity can be maintained, then you create scarcity by granting a monopoly to the first to innovate in a particular art. The fear of ‘freeloaders’ profiting using the borrowed innovations of others is a real fear, but only if you think that what is being used by the newcomer should not remain freely available for use and profit. The theory behind the justice of IP laws depends upon accepting the ethical stance that the fruits of one’s *intellectual* labor are somehow equivalent to the fruits of manual labor. For whereas we agree that if you build a house on a property you own, then you are entitled to occupy it unencumbered by anyone else’s claims to your property, do we likewise necessarily agree that if you come up with a good idea and express it for the world to see or hear, then others cannot use it freely? Many argue that the two situations are not ethically equivalent, as in the latter case, you are deprived of nothing by the use of your idea by another, whereas in the former, a squatter taking up residence in your living room does actually deprive you of the exclusive use of a property you built, enjoy, and justly own. Again, it is scarcity here that defines the ethics of the ownership and exclusivity. Simply put, IP is quite unlike other forms of property.¹⁹

The lack of scarcity of IP in a world prior to IP laws was apparent to Thomas Jefferson, an early proponent of limited IP laws, and a principle author of the first Patent Act. I quote him here at length because his insights are particularly important today, as we move into a new age of innovation:

It has been pretended by some, (and in England especially,) that inventors have a natural and exclusive right to their inventions, and not merely for their own lives, but inheritable to their heirs. But while it is a moot question whether the origin of any kind of property is derived from nature at all, it would be singular to admit a natural and even an hereditary right to inventors. It is agreed by those who have seriously considered the subject, that no individual has, of natural right, a separate property in an acre of land, for instance.

By a universal law, indeed, whatever, whether fixed or movable, belongs to all men equally and in common, is the property for the moment of him who occupies it, but when he relinquishes the occupation, the property goes with it. Stable ownership is the gift of social law, and is given late in the progress of society. It would be curious then, if an idea, the fugitive

fermentation of an individual brain, could, of natural right, be claimed in exclusive and stable property.

If nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself; but the moment it is divulged, it forces itself into the possession of every one, and the receiver cannot dispossess himself of it. Its peculiar character, too, is that no one possesses the less, because every other possesses the whole of it. He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me.

That ideas should freely spread from one to another over the globe, for the moral and mutual instruction of man, and improvement of his condition, seems to have been peculiarly and benevolently designed by nature, when she made them, like fire, expansible over all space, without lessening their density in any point, and like the air in which we breathe, move, and have our physical being, incapable of confinement or exclusive appropriation. Inventions then cannot, in nature, be a subject of property.

Society may give an exclusive right to the profits arising from them, as an encouragement to men to pursue ideas which may produce utility, but this may or may not be done, according to the will and convenience of the society, without claim or complaint from anybody. Accordingly, it is a fact, as far as I am informed, that England was, until we copied her, the only country on earth which ever, by a general law, gave a legal right to the exclusive use of an idea. In some other countries it is sometimes done, in a great case, and by a special and personal act, but, generally speaking, other nations have thought that these monopolies produce more embarrassment than advantage to society; and it may be observed that the nations which refuse monopolies of invention, are as fruitful as England in new and useful devices. (Thomas Jefferson, letter to Isaac McPherson, 13 August 1813)

Jefferson's insights are repeated today in the modern mantra that 'information wants to be free', something adopted more or less as a motto by the 'free software' movement and other opponents of IP. The problem we are now faced with, at the advent of the nanotech age, is whether IP poses more of a threat to the elimination of scarcity, increasingly rapid rates of innovation, and generally magnificent promises inherent in the transition from 'dumb' to programmable matter.

You Say You Want a Revolution?

In this book, I attempt to close an arc I began tracing more than a decade ago when I considered the problem posed by software patents, and proposed, among other things, a way to solve it. In *The Ontology of Cyberspace* I argued that the underlying ontology of expressions suggested that we could create a unified scheme of IP protection, and eliminate patents entirely. In *Who Owns You?* I expanded upon my thinking about IP as applied to biotech products, specifically the strange case of gene patents, and suggested that there were parts of the world that were simply, ethically not prone to ownership claims. I argued that there exist not only commons that we create where property claims would otherwise be legitimate (as in state parks, or other lands that are set aside purposely for public use), or what I called Commons by Choice. There is another type of commons, a Commons by Necessity, that includes parts of the universe that are simply unencloseable either as a matter of practical necessity (like sunlight) or as a matter of logical necessity (like laws or products of nature – e.g. genes). Now I am concerned with a promising new technology that many consider to be the ultimate convergence of all technologies. Nanotechnology and its currently working precursors (discussed in greater depth soon) are delocalizing manufacturing, opening up the process of innovation, and liberating intellectual capital all over the world. These technologies, if fully realized, will remake our economy, help to end scarcity of material goods, and incidentally reduce harms to the environment created by the necessity for both surplus and storage, but also shipping of goods. I will argue that if we continue to embrace traditional IP regimes, we will ultimately hinder the full realization of the promise of nanotech.

A clash is inevitable, however, because those who stand to profit by the IP industry, and who cling to the use of artificial scarcity to ensure profits,

are already threatened by a growing resistance to IP in current technologies. This is why we must now begin to choose which direction we will take, and how we can seize the ever more affordable tools of production, and work for institutional changes that will help us to truly liberate the full energy of human creativity.

Nina Paley, *Mimi & Eunice* (cc) creative commons