

The Basement Interviews

Preserving the Public Domain

Michael Hart, founder of Project Gutenberg, speaks to Richard Poynder

Published 9th March 2006¹

Michael Hart was born in Tacoma, Washington, USA, in 1947. His mother, who had earned three degrees before turning 18 and was a cryptanalyst during WWII, was business manager for "the chicest women's store" in the town; his father was an accountant.

When Hart was eleven his parents both decided to retrain as university professors, and the family moved to Urbana, Illinois, where Hart lives today. Even before entering high school Hart himself had regularly attended lectures at the University of Illinois. "I found nothing I couldn't understand," he says, "so long as jargon wasn't an issue."

When Hart went to University in his own right he completed a four-year "individual plan of study" program in two years, obtaining a degree in Human-Machine Interfaces. However, it was when — on 4th July 1971— that Hart was given an account on the University of Illinois mainframe that he found his true *métier*. Casting around for a worthwhile cause with which to use the computer time he had been given, and conscious that the computer was connected to the nascent Internet, he decided to type the US Declaration of Independence into the computer — a faux parchment copy of which he had just been given at his local grocery store to mark Independence Day.

Immediately seeing the potential of the network as a revolutionary new medium for distributing information, Hart was soon typing in entire books, including the Bible, all of Shakespeare, and *Alice in Wonderland*. Thus was born Project Gutenberg — a project that rapidly turned into an ambitious scheme to make electronic copies of 10,000 out-of-copyright books freely available on the Internet. Hart's mission: "to break down the bars of ignorance and illiteracy."

In retrospect Project Gutenberg was both prescient and revolutionary. In effect, Hart had become the first "information provider" twenty years before Tim Berners-Lee invented the Web, and at a time when there were, says Hart, just 100 people on the network. Indeed, what was to become the Internet was then viewed as little more than a powerful mechanism for crunching data — not a publishing platform.

¹ The interview took place in May 2005 (<http://poynder.blogspot.com/2006/03/basement-interviews.html>)

But for 17 years Hart had to plough a solitary furrow. Widely dismissed as "that crazy guy who wants to put Shakespeare in a computer" he had by 1997 input only 313 books. In 1998, however, he had a breakthrough. Linking up with the University of Illinois PC User Group, Hart set up a mailing list and started publicising his project. Soon he had recruited a team of volunteers to help, and by the end of that year 1,600 books had been keyed in. For his efforts Hart was honoured as one of the "Wired 25" in the November 1998 issue of *Wired* magazine.

Since then the number of volunteers has grown from tens, to hundreds, to thousands, and today Project Gutenberg offers over 17,000 e-texts, all of which can be freely downloaded in a wide variety of formats. In addition, there are now national Project Gutenbergs in Australia, Germany, Portugal, Canada and the Philippines, and plans are under way to create local projects in Africa, Asia, and other regions too.

New obstacles were to arise however: while copyright had always posed a challenge for Hart, the 1998 Sonny Bono Copyright Term Extension Act — extending US copyright by a further 20 years — removed one million potential eBooks from the public domain in one fell swoop. With copyright now averaging 95.5 years, and creators no longer needing to register their copyright, Hart began to fear that the public domain could disappear all together, undermining the *raison d'être* of what by then had become his life's mission.

When, therefore, opponents of the Sonny Bono Act decided to mount a legal challenge to the new law, Hart was the natural plaintiff. But Lawrence Lessig — the Stanford Law School professor who argued the case — refused to allow Hart to attach an appendix to the court documents expressing his personal views on the plundering of the public domain, so Hart walked away from the case, unprepared to be just a figurehead. His place was taken by Eric Eldred, and the case went all the way to the Supreme Court as *Eldred v. Ashcroft*, only to fall at the last hurdle.

However, with an unshakeable faith in the importance of what he is doing, and calculating that 9 million books will nevertheless enter the public domain by 2020, Hart set a new mission for Project Gutenberg: to make 10 million eBooks available in 100 languages.

For Hart the stakes are high, since he views Project Gutenberg as more than just the first and largest distributor of public domain eBooks. In addition, he argues, it is a primitive example of a "replicator" (a reference to a Star Trek machine envisaged as being capable of copying any inanimate matter by rearranging subatomic particles), and so therefore also a "lever to the Neo-Industrial Revolution."

Hart believes that in the future a powerful new breed of replicators — which will include 3D printers and nano assembly tools — will be able to produce and copy physical objects as easily and cheaply as it is currently possible to make and copy electronic books and other types of digital information. This, he says will enable a new "unlimited distribution" model that will eradicate scarcity and usher in a neo-industrial revolution, benefiting mankind as radically as did the first industrial revolution.

The danger, he cautions, is that to protect vested interests, governments and large corporations will seek to artificially maintain today's limited distribution mode, making the current controversy over the ownership of digital information and the Internet a mere dress rehearsal for a more profound struggle. In short, if the battle over intellectual property and the public domain is lost, says Hart, it will set an ominous precedent for the future.

Many are quick to characterise Hart as an eccentric. While rejecting that term, Hart agrees to "not being like anyone else I've ever met". Certainly he has an unusual take on the world — a characteristic that when Hart graduated led an assistant dean at the University of Illinois to comment: "Like many young men who are brilliant, there is always the danger of [Hart] being misunderstood as simply eccentric²." For that reason, he added, Hart should be "given the opportunity to pursue courses of action which might at first seem somewhat bizarre."

For its first 17 years Project Gutenberg was indeed viewed as a "somewhat bizarre" enterprise. Today, however, it is widely recognised as the first and most concerted attempt to preserve the public domain in the digital age.

Above all, however, Project Gutenberg is testimony to the determination of Hart, a man not easily deflected from his target. As his long-time friend Sue DeVries puts it: "Michael is absolutely brave, in the bone-deep sense, which makes him able to keep to a chosen path when others might find a task too daunting. He does not change his opinion or feelings with public opinion or outside pressure. He is also very funny, delights in all kinds of things, and has a great laugh."

As I discovered, Hart is also a challenging person to interview. Having a conversation with him is — as he himself puts it — like "herding cats". Nearly every question asked of him exploded into a series of related and more complex questions, and then bounced back to me. Since Hart's preferred mode of communication is e-mail, my inbox was soon throbbing with hundreds of messages; and when he eventually agreed to a telephone conversation he insisted that I turn off the tape recorder, leaving me desperately scribbling on a cascade of loose sheets of paper that constantly fell to the floor!

² <http://www.promo.net/hart/letters.html>

He also seems to be a man whose attention can quickly stray. After the hectic flurry of e-mails, and the phone conversation, I found it increasingly difficult to get Hart to respond to my requests for clarification of points he had made: his constantly active mind appeared to have wandered off — presumably to new schemes and ideas; and no doubt to more interesting questions from importunate journalists!

But eventually the interview was finished. What it reveals is a man whose whole life has clearly been dedicated to defending the public domain, but who is sometimes a little hard to fathom.



The interview begins...

RP: Were you a gifted child?

MH: My parents don't believe in such things and, until Sputnik went up, neither did the school system I was in. I was lucky: in 1958 my parents decided to go back to school and become professors — Mom chose math and Dad Shakespearean studies. So I got a great education at home on both fronts. I grew up in a house full of books, electronics (hi-fi) and artwork.

RP: What is your memory of school?

MH: School was just boring, except for baseball, marbles, dodge ball etc. School can be lots of fun if you don't care about grades. Grades are anathema to learning.

RP: Why anathema?

MH: When I was concerned about grades I never had a grasp of The Big Picture, it was all "Will it be on the exam?", whereas when I didn't care about the grades I studied the subject as a whole, remembered only what I thought was important, didn't worry about the exams so much, and got better grades anyway. When I was out for The Big Picture, I

would read the textbooks for fun, entertainment, *and* education, not just slave tasking, and the results were better in every way.

The more years I spent in school, the less attention I paid to grades, and in the end I finished college with straight A's in only two years. I had learned how to learn. Something they don't really teach in school. Young kids know how to learn, and then this is taken away from them as school goes on. I relearned it.

RP: You clearly also enjoyed regular activities like baseball and marbles — and indeed you were the champion of the first University of Illinois Frisbee Tournament! — but presumably you stood out from your peers? After all, you were still an undergraduate when you started Project Gutenberg — an initiative viewed today as both prescient and, in retrospect, really rather revolutionary?

MH: I don't think I stood out from my peers in any respect they would notice, certainly not in the way I dressed or acted. My interests were simply different from theirs, but most never realised it. I would never have been voted "mostly likely to" anything, since our interests didn't intersect.

The teachers didn't think all that much of me either, since I approached their subjects in ways that were not taught in the books. I did have one high school teacher who understood, and you can read her letter at: www.promo.net/hart.

RP: After interviewing you many journalists tend to describe you as an "eccentric", a description you appear to revel in. What kind of eccentric are you?

MH: Actually, I specifically avoid the term eccentric. However, I freely admit to not being like anyone else I've ever met. I don't think like other people do, otherwise you wouldn't be talking to me. I don't do the cookie cutter image thing. On the other hand, if you met me somewhere, I wouldn't stand out, other than that I will talk with nearly anyone.

RP: So how would you describe yourself?

MH: Idealistic, utilitarian, workaholic, innovative, independent, intelligent, adventurous, gregarious. In short, I am an idealist, and a person who puts his thoughts and ideals into action; a visionary, if you will. Most people still do not understand as much about eBooks as I understood in that first moment when I started Project Gutenberg thirty-five years ago.

A light went on

RP: Well, it was certainly not what you would expect from the average undergraduate. I assume it was no accident either that the first document you made available was the US Declaration of Independence?

MH: Actually it was pretty serendipitous! Project Gutenberg was just one of those great combinations of luck and being the right person in the right place at the right time.

RP: How do you mean?

MH: We were coming up to the 200th anniversary of the War of Insurrection. On the night I started Project Gutenberg I had been out later than usual watching the 4th of July fireworks. On my way back I decided to get some food and stopped at the grocery store, where a faux parchment copy of the US Declaration of Independence was given to me with my dinner.

As luck would further have it, my brother's best friend was the University's mainframe operator, and that very night I had been given my first computer account — along with \$100,000 of the "play money" required to run it. So I was trying to figure an interesting and worthwhile way to use the account.

After spending an hour or two trying to think of something, I decided to eat, and as the faux parchment fell out with the groceries a light went on over my head. As a result, on the night of July 4th, 1971, Project Gutenberg came into existence, as I sat down and typed in that historical document.

RP: This was long before the Web. Making it available on the Internet in 1971³ presumably meant loading it directly on to the university mainframe. How exactly did you "publish" the document?

MH: I used a Teletype⁴ machine (possibly the one I still have in my basement) with paper tape and all that. My initial plan was to email it — after all, just 100 people then inhabited the Net⁵. But I was told that I would crash the entire system if I did. It's hard for

³ Technically the Internet did not exist in 1971. It developed out of a network called ARPAnet. ARPAnet had been created by the US Advanced Research Projects Agency (ARPA) in 1969. One of ARPA's goals was to connect mainframe computers at different universities around the country so that they would be able to communicate using a common language and a common protocol. In the 1980s the ARPAnet gradually merged into the Internet, and was itself closed in 1994.

http://en.wikipedia.org/wiki/History_of_the_Internet

⁴ A teletype (teleprinter, teletypewriter, or TTY) is a now largely obsolete electro-mechanical typewriter which can be used to communicate typed messages from point to point through a simple electrical communications channel, often just a pair of wires. <http://en.wikipedia.org/wiki/Teletype>

⁵ How accurately can we estimate the number of users on the Internet? Certainly it is not a precise science, particularly when looking back 35 years. Some argue that a more accurate measure is the number of hosts (computer systems with a registered IP address). Hobbes' Internet Timeline indicates that there were 23 hosts connected to ARPAnet in 1971 (<http://www.zakon.org/robert/internet/timeline>). The Internet Systems Consortium (ISC) estimates that there are now nearly 400 million Internet hosts. To return to users, last year the *Computer Industry Almanac* estimated there are now one billion users on the Internet, a figure it

people to realise just how primitive the Net was then — we think nothing of sending a five gigabyte file to thousands of people, but one five kilobyte file to 100 people would have shut it all down.

So I put a message on what later became comp.gen, a Usenet newsgroup⁶, and kept the files in the school's tape farm. People would send an email if they wanted to see it and someone in the computer lab would load the tape, usually my best friend. Then we would tell them where it was and how to download it.

RP: A small start, but presumably your choice of name for the project implied high ambitions: Johannes Gutenberg⁷ was the 15th Century printer who invented the moveable type printing press; since everything had to be handwritten before then he enabled a huge revolution in the way information could be disseminated.

MH: Even the greatest journeys start with but a single step. The ramifications of the Gutenberg Press are far more profound than I have ever heard anyone even try to mention. The social, economic, and intellectual changes were all revolutionary.

RP: Right. Amongst other things, Gutenberg enabled translations of the Bible to be made widely available — a development that later led to Luther nailing his 95 theses to the Wittenberg church door, and so to the protestant reformation.

MH: "Translation" is one of the key words here. The church did not want the masses to be able to read the Bible. Gutenberg printed it in a way they could read it, by themselves, in their own language, rather than the long-dead and secret Latin used by the church.

RP: Unlike Gutenberg you didn't invent the technology but....

MH: Actually, Gutenberg didn't invent the printing press; he just added one element: moveable type. I didn't invent the Internet; I just added one element: books that can be copied at the speed of light. Like Gutenberg I made books so much easier for the masses to get, and so much cheaper.

RP: The point I wanted to make is that you were the first person to see the potential of the Internet as a new information medium, rather than a tool for crunching data. In

predicts will rise to two billion by 2011 (<http://www.c-i-a.com/pr0106.htm>).
<http://www.isc.org/index.pl?ops/ds/host-count-history.php>

⁶ Usenet is a distributed Internet discussion system that was conceived by Duke University graduate students Tom Truscott and Jim Ellis in 1979. Users read and post email-like messages to a number of distributed categorised newsgroups. The medium is distributed among a large number of servers, which store and forward messages to one another. Individual users download and post messages to a single server, usually operated by their ISP or university, and the servers exchange the messages between each other.
<http://en.wikipedia.org/wiki/Usenet>

⁷ Johannes Gensfleisch zur Laden zum Gutenberg was a German metal-worker and inventor who achieved fame for his contributions to the technology of printing during the middle of the 15th Century, including a type metal alloy and oil-based inks, a mould for casting type accurately, and a new kind of printing press based on presses used in wine-making. http://en.wikipedia.org/wiki/Johann_Gutenberg

*effect, you became the first Internet information provider twenty years before Tim Berners-Lee created the Web*⁸.

MH: I was certainly the first person to say that computers and the Internet would *not* to be used for computing, but for the dissemination of general information. But it took a long time for the idea to catch on, and no one really paid any attention for some time. In fact, it was 17 years before I received any comment other than I was "that crazy guy who wants to put Shakespeare in a computer". But I immediately saw what the Net would be like 35 years down the line, except that I was not as concerned with multimedia as everyone else was.

RP: *What was also revolutionary about Project Gutenberg is that it became one of the first examples of the network being used to enable a group of geographically dispersed volunteers to collaborate on a common task — a model later famously exploited by the Open Source movement. How did the collaboration start?*

MH: As I say, not for a long time since there was so little interest in what I was doing. The big change came when I went to a few demos by Apple and IBM on the University of Illinois campus, and met the people who were about to start the University of Illinois PC User Group. This was around 1988-89.

At the first PCUG meeting I was the only one who actually brought a computer, so by default I set up the [the group's] mailing list and wrote the first newsletter. Mark Zinzow⁹ was also at that first meeting and we became friends and started one of the Tuesday lunches I still go to 15-16 years later.

Mark was a huge help in getting things rolling for the Project Gutenberg mirror sites, list servers, mail servers, etc. There is no way to thank him enough. After a while volunteers started coming from all over. By then the Net had passed a quarter of a million people and it was possible to find someone interested in anything, so getting volunteers became much easier.

RP: *How many volunteers are there now?*

MH: I would estimate that we now have between 10,000 to 20,000 volunteers.¹⁰

⁸ Tim Berners-Lee developed the concept and the protocols of the World Wide Web in 1989 (<http://www.w3.org/People/Berners-Lee>), but he did not release them on the Internet until 1991.

⁹ Mark Zinzow is a senior research programmer at the University of Illinois

¹⁰ Initially the text of an eBook was physically typed in, and one person or a small group of like-minded individuals undertook the process of proofreading and preparing it. In late 2000, however, Distributed Proofreaders (<http://www.pgdp.net/c/default.php>) was launched to support the development of Project Gutenberg, and is now the main source of PG eBooks. Project managers scan the books and the images are run through optical character recognition (OCR) software. OCR software, however, is presently far from perfect, and a large number of errors often appear in the resulting text. To deal with this, individual pages are made available to volunteer proofreaders via a web-based interface. This displays the original page's image and the recognised text side-by-side. Using only a web browser, therefore, volunteers can download a single page of OCR text and its matching page image, make changes to the text, and save it. After two such passes the pages are returned to a project manager who fixes any problems noted by the proofreaders

RP: The original mission of Project Gutenberg was to create 10,000 eBooks. You achieved that in 2000. What's next?

MH: Right. Contrary to all the pundits' prolific predictions over the years, Project Gutenberg has not only not gone under, but has continued its growth on a curve in excess of Moore's Law¹¹.

The initial aim was to offer 10,000 eBooks and get the average eBook to 100 million readers, or 1.5% of the world. We now have 17,000 eBooks and the current mission is to create 1,000,000 eBooks and get the average eBook to 1 billion readers, or around 15% of the world. In total, we now provide eBooks in 42 languages.

RP: What's the end game?

MH: The future mission is to create 10,000,000 eBooks and translate them into 100 different languages.

RP: Effectively, you plan to make every book in the public domain available as an e-text?

MH: Right. Once Project Gutenberg has a million items to offer, it should be an easier task to add the remaining 9 million items that it is estimated will become available in the public domain between 2010 and 2020. Then in its final stages Project Gutenberg will focus on collecting materials from all 100 languages, and disseminating them in other languages. So the eventual aim is to be able to offer 10 million eBooks in 100 languages to as many readers as possible.

Legal threats

RP: Project Gutenberg is restricted to inputting pre-1923 texts due to copyright law right?

MH: Yes, although around 2% of the e-texts are still in copyright. Believe it or not, with books still in copyright, the authors or copyright holders usually approach us.

In addition, most copyrights were never renewed as per US copyright law. In fact, 90% of them were never renewed, so once we find out all of the renewal data, there will be more eBooks.

and then submits the e-text to Project Gutenberg. This process effectively distributes the time-consuming error correction process in an analogous way to distributed computing. It also means that it is difficult to estimate the number of volunteers at any one time.

¹¹ Moore's law is about the empirical observation, that at our rate of technological development, the complexity of an integrated circuit, with respect to minimum component cost, will double in about 18 months. http://en.wikipedia.org/wiki/Moore's_law

RP: Can you expand on that? I understand US law was harmonised with Europe, removing the need to register or renew copyright. Are you saying that copyrights not renewed prior to the change in US law have fallen into the public domain regardless of when they were written?

MH: Up to a certain point, yes. More recent copyrights were grandfathered in to the new term, but older ones were not. So it's not always easy to figure out, even for the copyright experts.¹²

RP: What issues does copyright raise for the work of Project Gutenberg?

MH: The biggest problem is the time it takes to do the necessary copyright research before making an e-text. The next biggest is dealing with the constant threat of lawsuits, which come in every year.

RP: You get a lot of legal threats?

MH: We do. Most are just obvious smokescreens from the start — lawyerly types who expect us just to fall over in shock. I think they know they are wrong to start with, and just figure we will all crumple when they blow smoke in our direction.

However, we have a number of very highly placed copyright lawyers, and heads of school law departments, who help us. Once those threatening to sue us realise we know more about copyright than they do, and have done our research correctly, they all vanish in a puff of smoke.

But the biggest copyright problem we ever faced was after we spent nearly the entire 1980s working on an edition of Shakespeare that was expected to go into the public domain, but failed to do so when the copyright laws were changed¹³.

¹² As Hart indicates, these are complex matters. Today US copyrights last for seventy years after the death of an author, or seventy-five to ninety-five years in the case of works of corporate authorship and works first published before January 1, 1978. Until 1992 a copyright renewal request had to be filed with the Library of Congress Copyright Office in a work's 28th year to have copyright extended. When the need for renewal was eliminated, works that had already entered the public domain by non-renewal did not regain copyright protection. So works published before 1964 that were not renewed are public domain. Consequently some material from as recently as 1963 has entered the public domain but some as old as 1923 remains copyrighted — if renewals were filed. In addition, when The Copyright Term Extension Act of 1998 — alternatively known as the Sonny Bono Copyright Term Extension Act — extended copyright terms in the United States by 20 years the act also affected copyright terms for copyrighted works published prior to January 1, 1978, increasing their term of protection by 20 years as well. Under this act, no additional works made in 1923 or afterwards that were still copyrighted in 1998 will enter the public domain until 2019. Effectively the Sonny Bono Act 'froze' the advancement date of the public domain in the United States for works covered by the older fixed term copyright rules. Today all works published in the United States before 1923 are in the public domain, but works created before 1978 but not published until recently may still be protected until 2047.

. http://en.wikipedia.org/wiki/Sonny_Bono_Copyright_Term_Extension_Act

¹³ i.e. The Sonny Bono Copyright Term Extension Act

RP: It seems puzzling that you could be stopped from publishing a book of plays by an author who died 390 years ago. How can this happen?

MH: Because a publisher can obtain a new copyright for a new edition of Shakespeare by combining pieces from previous editions.

RP: The change in the law you refer to is the 1998 Sonny Bono Copyright Term Extension Act, which extended the length of copyright in the US by 20 years¹⁴?

MH: Yes. As a result, US copyright now averages 95.5 years, and the public domain has been shrunk by one million books, not to mention the loss of magazines, newspapers, movies, and music.

RP: Certainly the Sonny Bono Act has been widely criticised, and opponents challenged its legality all the way to the Supreme Court. You were the original plaintiff in that legal challenge, but Eric Eldred subsequently replaced you. As I understand it, this was because you and Lawrence Lessig — the Stanford Law School professor who argued the case¹⁵ — disagreed over how to conduct the matter. Specifically, you wanted to attach to the court papers a personal statement outlining your views on the greed of copyright holders. Is that correct?

MH: All I asked was that I be able to write an appendix containing my views. I simply wanted to demonstrate the magnitude of the changes, and point out that the Act had voided a social contract and replaced it with a new one. I tried not to be inflammatory, but it's hard not to be when the public domain is being raped and pillaged right in front of you.

RP: Lessig, however, refused your request. Wired magazine¹⁶ quoted him as saying, "Our view was that populist appeals are great, but you've got to frame a constitutional argument." In response, added Wired, you said: "Enough — you can't use my name".

MH: That's what I said, but not in that context. Initially Lessig said, "You're the client, you're the boss." But after I had spent five months writing draft after draft for Lessig he admitted that he would never include anything I said. That was breach of contract, and I was not willing to be a figurehead.

¹⁴ The act was named after the late Congressman Sonny Bono (who was killed in a ski accident more than nine months before the act became law), who had favoured this position as a songwriter and filmmaker even prior to his entry into politics. <http://www.imdb.com/name/nm0095122>

¹⁵ Lawrence Lessig is an American academic. He is currently professor of law at Stanford Law School and founder of its Center for Internet and Society. He is best known as a proponent of reduced legal restrictions on copyright, trademark and radio frequency spectrum, particularly in technology applications. He is founder and chairman of the Creative Commons and a board member of the Electronic Frontier Foundation. http://en.wikipedia.org/wiki/Lawrence_Lessig

¹⁶ *Lawrence Lessig's Supreme Showdown*, Steven Levy, Wired, 10.10, October 2002, http://www.wired.com/wired/archive/10.10/lessig_pr.html

RP: *As it turned out, the Supreme Court threw the case out, and Lessig subsequently conceded that he had overplayed the constitutional argument. Writing in Legal Affairs¹⁷ in 2004, he said he now realises he would only have persuaded the justices to vote against the media companies if he had shown that "dramatic harm were being done to free speech and free culture." But do you think your approach would have been any more successful?*

MH: He was wrong, and that's why we parted company. And there was no reason why we couldn't have used both approaches. However, I knew he wasn't actually trying to win, because I read every word of every draft.

RP: *Can you explain what you mean?*

MH: Going to SCOTUS¹⁸ is not like the Super Bowl, where you only get the jewellery to wear if you win. Larry got his sterling silver quill pen even though he lost. I wanted to win: he was there just to make an appearance. I told Larry I would never forgive him until after he gave me his pen.

I think Larry hurt the public domain as much as anyone ever did, including the Sonny Bono Act. Perhaps even worse, since SCOTUS won't hear another such case for a century, when 99.99% of all copyrights ever given will still be in force, and it will be too much to change back easily — so we won't have much of a public domain. You will not believe the numbers, even if you understand them.

RP: *Try me. How much is being lost to the public domain as a result of the Sonny Bono Act?*

MH: The answer is about \$10 trillion. This is based on a calculation of just one cent per book per lifetime, and assumes only about 15% of the world are readers, and only one million books are being lost to the public domain.

RP: *You say your current target for Project Gutenberg assumes 9 million items coming into the public domain by 2020. Over the next century, however, you anticipate a situation in which 99.99% of all copyrights will still be in force. In effect, you are talking about the gradual death of the public domain. Can you talk me through that?*

MH: Sure. If you take the US situation, for instance, and consider that a century ago copyrighted information was doubling roughly every 15 years, and that US copyright was around 15 years in duration, then you can easily compute that at any given time about 50% of all copyrights had expired — giving an even balance between the amount of copyrighted information and the amount of information in the public domain.

¹⁷ *How I lost the Big One*, Lawrence Lessig, March/April 2004, http://www.legalaffairs.org/issues/March-April-2004/story_lessig_marapr04.html

¹⁸ The Supreme Court of the United States

Now, ignore for the moment the impact of any copyright extensions, and consider that the doubling rate for copyrighted information has been increasing very quickly over the last century — a process that by itself increases the percentage of information under copyright, and correspondingly decreases the percentage of public domain information.

So if you assume that US copyright had stayed the same, if information were to double in around 7.5 years, then 75% of copyrights would remain unexpired at any given time. If you then assume that it is doubling in around 3.25 years, then 87.5% of copyrights remain unexpired at any given time.

So moving from a situation where around 50% of information is copyrighted to a figure of 87.5% is a world-shaking event. In fact, however, the current consensus amongst copyright authorities is that copyrightable information is now doubling at least as fast as every 3.25 years, with some estimating that it is doubling at around every 1.8 years.

Whatever the precise figure, it is widely agreed that information is growing at a very substantial rate. As a consequence, even with a constant copyright term, an ever-growing portion of all copyrightable information currently is under copyright at any given time.

***RP:** OK, you tested me with the numbers, and I'm going to trust your math! But what you are talking about, surely, is a product of an information explosion, not of any changes to the law?*

MH: Right. But what we face is a situation in which the rise in the percentage of copyrighted information due to growth is being compounded by the constant extension of copyright terms. Put these two elements together and you end up with a situation in which well over 90% of all copyrights ever granted are still in force.

In other words, if you combine the information explosion with the fact that the average copyright term has risen from about 15 years to around 95 years then the public domain — which was about 50% of everything every written a century ago — will fall to around 0.00001% or less just a century into the future.

***RP:** In reality, of course, the creators of much of this material are probably happy to have their works enter the public domain. The problem lies in the difficulty of tracing copyright owners in order to confirm that they are, in fact, the owners — a phenomenon often referred to as the "orphan works" problem¹⁹. If copyright owners were again required to register their copyright the problem of the shrinking public domain might be*

¹⁹ Orphaned works are, broadly speaking, any copyrighted works where the rights holder is hard to find. Because the cost of finding the owner is so high, creators cannot build on orphan works, even when they would be willing to pay to use them. In many cases the works were abandoned because they no longer produced any income. In most cases, rights holders, once found, are delighted to have their work used. Some countries have considered a compulsory license scheme for orphaned works; Canada has enacted one. http://en.wikipedia.org/wiki/Orphan_works

*partially addressed. I understand the US Copyright Office is currently looking at possible solutions like that?*²⁰

MH: That's not an issue I have been following closely. But believe me, they will never let it happen.

Bear in mind that the Sonny Bono Act was not the only time copyright law was changed in the last century. There have now been three copyright extensions in the US in the last 100 years. In 1909 it was done to counteract the steam and electric presses; in 1976 it was done to counteract the Xerox machine; and in 1998 it was done to counteract the Internet.

As I say, the average length of copyright is now 95.5 years, which is in stark contradiction to the original idea of a 14-year copyright as created by the Statute of Anne in 1709-10²¹. The problem is that if you keep changing the copyright term before it expires, how can anyone believe in copyright?

RP: *You clearly see a larger agenda here. In a previous interview²² you said: "The Internet could be the greatest educational tool of all history, but they are trying to pass laws to keep all the great books of the last century off the Internet."*

MH: Right. Just to clarify: I meant the most recent century, not the one before. But it just seems obvious to me that they would never allow that kind of power into the hands of the public, no matter how much they hypocritically talk about "universal education."

RP: *Who are "they"? Do you mean governments, or large media companies?*

MH: Both: the media lobbies wholly controlled the Sonny Bono Act. You probably don't realise that the passage of that law was hidden behind one of the biggest smokescreens of all time: the impeachment of President Clinton. The Republicans knew they couldn't win, but they passed the copyright law in the same 24 hours as the articles of impeachment.

And this is not just about profits. Do you think the best sellers from 70 years ago are really going to make more than 1% additional profits from 20 more years of government and WIPO²³-sponsored monopoly? No, the aim is to stop alternatives to new "pay to read" publishing models, to kill the public domain, and to remove us from control of our histories.

Project Gutenberg, by the way, is not just a distributor of eBooks; it's a lever for "The Neo-Industrial Revolution."

²⁰ The US copyright office recently submitted its 200-page report on orphan works to the US Senate. This report, along with much supplementary material, is now available online at: <http://www.copyright.gov/orphan>

²¹ The Statute of Anne was the first copyright law in Britain, enacted in 1709 and entering into force on April 10, 1710. It is generally considered to be the first fully-fledged copyright law. It is named for Queen Anne, during whose reign it was enacted. http://en.wikipedia.org/wiki/Statute_of_Anne

²² *Netpanel, 1997*, <http://www.netpanel.com/articles/misc/mikehart.htm>

²³ World Intellectual Property Organisation

The Neo-Industrial Revolution

RP: What is the Neo-Industrial Revolution?

MH: In 1455 Johannes Gutenberg introduced the idea of moveable type, and thus interchangeable parts. In doing so he started the industrial revolution. Project Gutenberg is the first primitive example of a replicator, and so the start of the Neo-Industrial Revolution, which is tied to the computer in the way the Industrial Revolution, was tied to interchangeable parts.

RP: This term replicator comes from Star Trek I think²⁴. Replicators were machines capable of converting energy into matter and vice versa, and so able to copy any inanimate matter by rearranging subatomic particles.

MH: Right. And today you can digitise something and put it anywhere on the Internet — be it an eBook, a song, or a picture of The Mona Lisa. And a minute after it is posted, anyone on the Net can, for the rest of history, download it.

So we are moving from a "limited distribution" model to an "unlimited distribution" model. As we enter the Neo-Industrial Revolution we will see more and more things being digitally copied. Eventually practically everything will be digitisable and copyable, and so part of the replicator system.

RP: Digitisation is about copying information, not inanimate matter, isn't it?

MH: Are you aware how many three dimensional things can be made via the new technology of "stereolithography"?²⁵ Are you aware that 3D "printers"²⁶ have been around for two decades?

RP: Well I've seen you elsewhere²⁷ claim that 30 years from now we might be able to "print out" a wide variety of physical materials. I recall a particular quote in which you

²⁴ [http://en.wikipedia.org/wiki/Replicator_\(Star_Trek\)](http://en.wikipedia.org/wiki/Replicator_(Star_Trek))

²⁵ Stereolithography is one of the more commonly used rapid manufacturing and rapid prototyping technologies. It involves building plastic parts a layer at a time by tracing a laser beam on the surface of a vat of liquid photopolymer. The photopolymer is solidified by the laser light. Once one layer is completely traced, it is lowered a small distance into the liquid and a subsequent layer is traced, adhering to the previous layer. After many such layers are traced, a complete 3D model is formed.

<http://en.wikipedia.org/wiki/Stereolithography>

²⁶ Three-dimensional printing (more commonly referred to the abbreviation 3DP) is a type of rapid prototyping system developed by MIT in the late 1980s. Using an adapted inkjet printing system, layers of a fine powder (either corn flour or plaster) are selectively bonded by "printing" a water-based adhesive from the inkjet print head in the shape of each cross-section as determined by a CAD (computer aided design) file. Alternately, these machines feed liquids into individual jetting heads, which squirt tiny droplets as they are scanned to form a layer of the model. The liquid hardens after being deposited. Once a single layer has been deposited, a milling head may be employed to ensure uniform thickness before the next layer is deposited. http://en.wikipedia.org/wiki/3D_printer

²⁷ United Press International, 2002, <http://www.upi.com/view.cfm?StoryID=28052002-120328-3546r>

said that it would be possible to "print" a car, and then get into that car and drive it away. This is all rather far fetched, isn't it?

MH: Think about it this way: initially computers had only 2 colours, then 4 colours, 16 colours, 32 colours, 128 colours, 256 colours, and now all seem to have a minimum of 16 million colours.

Printers can follow the same evolution. We already have the Crayon-Jet Printer, which is fed by simple crayons available at stores all over the world extremely inexpensively.

RP: *How do these printers work?*

MH: You tear the paper off the crayons and feed them into hoppers in the printer. The tip of the crayon is heated and jetted on to a paper or cardboard surface, layer after layer, and eventually builds up to the desired 3D surface. So a printer that "prints" in a 100th inch layer would take 100 passes to "print" a map for a relief scale of one inch, or whatever you programmed.

So you can imagine that in the educational context you could print out all the planets for an astronomy class, and make them to exactly the right scale and colour. You could print out famous statues for art class, famous paintings, sculptures, or city blocks, complete with trees, and cars, and people. The possibilities are endless!

RP: *That's still a long stretch from printing a car and driving it away.*

MH: If the materials used with 3D printers can double as fast as the number of colours for computers, then we will have 3D printers that can print with millions of materials in just 20 years. In addition, nanotech "printers" will be creating more and more 3D materials, which will give us even more possibilities.

RP: *Well, I realise that many believe "nano-printing" will enable the very cheap mass production of nano-devices²⁸. But what you seem to be implying is nano assembly, in which objects would be built atom by atom. This is all somewhat theoretical today, isn't it?*

MH: Not only is it possible, but I think it is probable — even within the lifetime of people living now. And it will allow near costless construction, and unlimited distribution.

RP: *What you are saying then is that physical objects will become more like digital information. That is, you will be able to copy them as effortlessly and limitlessly as it is now possible to copy an eBook, or, indeed, as simply as Napster copied songs. That is what you mean by unlimited distribution?*

MH: Yes.

²⁸ http://www.eurekaalert.org/pub_releases/2005-06/miot-mnc060705.php

RP: So where today the cost of producing goods lies primarily in the raw materials, and the cost of manufacturing, in the world you envisage most of the cost, or value, would lie in simply knowing how to make them, or having access to the necessary tools?

MH: Right. However, given that civilisation as we know it has been based on a structure of limited distribution, the idea of unlimited distribution is going to be fought tooth and nail.

RP: And the way to fight it would be to use the intellectual property system to limit the right to make the goods, or use the tools to make them, and so create artificial scarcity and limited distribution?

MH: Exactly. In a virtual world it's less a question of ownership in the traditional sense, but who has control. So the future will be a choice between unlimited distribution and an end to scarcity; or limited distribution and a world in which artificial controls are placed on the potential for unlimited distribution.

RP: If you are right, then today's disputes over the ownership of information and intellectual property are simply a dress rehearsal for a more profound struggle.

MH: Absolutely. It will surely lead to a war between a proprietary *rentier*²⁹ economy, and those who support the Neo-Industrial Revolution, which promises a world in which everyone can have everything.

RP: The development of the Internet is clearly leading to a growing conflict between proprietary interests and supporters of a more communal approach. Do you see Project Gutenberg as part of a wider movement for open, or free, information?

MH: No, I don't see Project Gutenberg as part of any movement; we just give away eBooks; if other people latch on to some of our principles, or principal ideas, that's just great.

RP: Nevertheless, the aims of the various open and free movements that have developed seem very similar to those that we are discussing in relation to Project Gutenberg. I'm thinking about the Free and Open Source Software movements, the Open Access movement, Open Spectrum, and Creative Commons. Do you take an interest in these other movements?

MH: I am aware of them to some degree, but not influenced by them, as they are much more political than I have any intention of being.

RP: You don't see Project Gutenberg as a political cause then?

²⁹ A person who lives on income from property or investments

MH: No. I am apolitical, to the point of being anti-political. I don't believe in the modern political process at all.

RP: The mission statement of Project Gutenberg is to "break down the bars of ignorance and illiteracy." Does that not imply a wish to empower people politically by giving them free books?

MH: Breaking down the bars of ignorance and illiteracy is only political if someone else is trying to keep them up! A political statement tries to set policy; I don't set policy. I just try to lead by example. eBooks are revolutionary in the non-political sense, and yet they could change the world as much as the Gutenberg Press did.

RP: So what is the aim of Project Gutenberg?

MH: Project Gutenberg is about education and literacy, and providing a level playing field. I don't try to change people; I try to change the information infrastructure. We make eBooks and give them away. We don't choose which books or where they should go. What people do with them is totally up to them.

A dozen other formats

RP: Has the disappointing take-up of dedicated eBook readers led you to put a greater emphasis on other platforms?

MH: No. We don't put any emphasis on platforms. We try to make our eBooks work on as many combinations of hardware and software as possible. So you'd have to find a very specialised computer system to find one that would not read nearly all Project Gutenberg eBooks.

RP: You have also made a point of always ensuring that — whatever other formats might be provided — there is always a plain ASCII file available, right?

MH: We have. Plain ASCII³⁰ is still what people seem to use most, along with HTML — so most files are available in HTML and plain text. "Read-aloud programs"³¹ work best with plain text for instance, as does changing the font size. Many visually challenged people love the Project Gutenberg eBooks for this reason: they have control over how the book looks, and how it works in text-to-speech programs; control they don't have with other kinds of eBooks.

³⁰ ASCII (American Standard Code for Information Interchange), is a character encoding based on the English alphabet. ASCII codes represent text in computers, communications equipment, and other devices that work with text. <http://en.wikipedia.org/wiki/ASCII>

³¹ Read-aloud software includes text-to-speech reading programs that can read aloud text that has been saved using a regular word processor, highlighting, and simultaneously reading aloud each word. Many are designed specifically for people with a learning disability, or for people who struggle to read text. <http://www.abilityhub.com/ecolumn/2002/oct1.htm>

But we also have a dozen other formats available, including MobiPocket³² and Palm³³ files, Word files, PDF, and a variety of music, image, and markup formats. This means that people can use notebook computers, PDAs, and now even cell phones. Just about a week ago I got an email from Africa telling me how people were both reading and listening to eBooks on their cell phones there. It's hard to believe, but when you are out on the Serengeti plains³⁴, you still get cell phone service.

The point is that there are billions of cell phones out there; billions of laptops, and billions of desktops. There are not even millions of eBook readers. The results are obvious — percentage wise, no one is reading on dedicated readers.

RP: And you now also offer DVDs on which the books are available not as text, but with each page captured as an image. This includes "The Million Dollar DVD" and "The Second Million Dollar DVD". Can you explain the titles?

MH: The message is that the average hardback book this year will be over \$40. So 30,000 eBooks on a single plain DVD contains a million dollars worth of books — about the same number as the average US public library. We, however, will offer this free of charge, as we do with all our products.

RP: And the range of information you offer is growing too. You host human genome data, for instance, and there is also something called the Gutenberg radio project.

MH: We have three editions of the Human Genome Project online³⁵. G-Radio³⁶ is basically just audio eBooks. We will also be adding more multimedia eBooks, paintings, sculptures, music, movies, etc. In addition, there are newspapers, magazines, and the various other types of media.

A different kind of power

RP: How much hands-on work do you still do for Project Gutenberg?

MH: Sadly, I've been relegated to coordination and management now, although I still write the weekly newsletter. Nevertheless, it is still more than a full-time job for me.

³² <http://www.mobipocket.com/en/HomePage/default.asp>

³³ <http://ebooks.palm.com>

³⁴ Serengeti is a region of grasslands and woodlands in Africa shared between the countries of Tanzania in the south and Kenya in the north. The whole region is spread over around thirty thousand square kilometres. Eighty percent of this region lies in Tanzania. <http://en.wikipedia.org/wiki/Serengeti>

³⁵ e.g. <http://www.gutenberg.org/etext/11799>

³⁶ <http://www.gutenberg.org/audio>

RP: Do you get a salary from the Project Gutenberg Literary Archive Foundation, which was founded, I believe, in 2000?

MH: No. We don't attract enough funding for that.

RP: So what do you live on today?

MH: It's been two years since my last pay check, but if you save all your salary when you do get one, \$100,000 will go 10 years with no salary, at \$10,000 a year.

RP: And when the savings run out?

MH: I don't expect to let them run out. I'll either be out of commission, or have brought in more money by then. The fact is that I have just never dealt with money all that much in my whole life. Money was never a big thing for me. I'm into a different kind of power. Since I never spend much, and save everything, money isn't an issue at all.

I know that sounds odd to most people, but I just never bought into the money system all that much. I never spent it when I got it. It's all a matter of perspective; most people spend the vast majority of their money on things I just don't care about. The only time I notice it is if I think of buying a new house.

RP: So you lead a somewhat Spartan lifestyle?

MH: Well, it's hard for me to spend \$10 on dinner; the average is well under \$5. I have no cable or cell phone. I ride a bicycle most of the time, and I only buy a car when I am faced with a long trip and the cost is too high; then I buy a cheap one, make the trip, and then make it last a couple more years.

I also wear garage sale clothes; in fact I live just about totally on garage sales³⁷, and living in a university town, I find there are sometimes over 200 garage sales on a single day. Garage sales also offer a great chance to just meet people all over town, and paw through their stuff, and talk about everything. I also go to them wherever I travel: it helps me find out about the people who live there, and bring back souvenirs, etc.

RP: Let's leave it on that note. Thank you for talking to me.

MH: You are most welcome.

© 2006 Richard Poynder

³⁷ A garage sale (also known as a yard sale, rummage sale, or jumble sale) is an informal, irregularly scheduled marketplace of used household goods, typically sold by one or at most a few families. In some communities there are designated days every year in which "block sales" are allowed, so that people don't have to get the required permits or collect sales tax. http://en.wikipedia.org/wiki/Garage_sale

This interview is distributed under the terms of the Creative Commons Attribution-Non-commercial-No-Derivatives Licence (<http://creativecommons.org/licenses/by-nc-nd/2.0/>). This permits you to copy and distribute it as you wish, so long as you credit me as the author, do not alter or transform the text, and do not use it for any commercial purpose.

If you would like to republish the article on a commercial basis, or have any comments on it, please email me at richard.poynder@journalist.co.uk.

Please note that while I make this interview freely available to all, I am a freelance journalist by profession, and so make my living from writing. To assist me to continue making my work available in this way I invite anyone who reads this interview to make a voluntary contribution. I have in mind a figure of \$8, but whatever anyone felt inspired to contribute would be fine. This can be done quite simply by [sending a payment](#) to my PayPal account quoting the email address aotg20@dsl.pipex.com.

I would like to acknowledge the help of the [Open Society Institute](#), which provided a small upfront grant to enable me to get started on *The Basement Interviews* project.

Further information about *The Basement Interviews* can be found at my blog [Open and Shut?](#)