On Crime and Access to Knowledge

Carl Malamud

ON CRIME
AND
ACCESS TO KNOWLEDGE

AARON'S ARMY

ON CRIME AND ACCESS TO KNOWLEDGE

CARL MALAMUD

1   Do not think for a moment that Aaron's work on JSTOR was the random act of a lone hacker, some kind of crazy, spur-of-the-moment bulk download.

2   JSTOR had long come in for withering criticism from the net. Larry Lessig called JSTOR a moral outrage in a talk and I suppose I have to confess he was quoting me. We weren't the only ones fanning those flames.

3   Sequestering knowledge behind pay walls—making scientific journals only available to a few kids fortunate enough to be at fancy universities and charging $20 an article for the remaining 99% of us—was a festering wound. It offended many people.

4   It embarrassed many who wrote those articles that their work had become somebody's profit margin, a members-only country club of knowledge.

5   Many of us helped fan those flames. Many of us feel guilty today for fanning those flames.

6   But JSTOR was just one of many battles. They tried to paint Aaron as some kind of lone-wolf hacker, a young terrorist who went on a crazy IP killing spree that caused $92 million in damages.

7   Aaron wasn't a lone wolf, he was part of an army, and I had the honor of serving with him for a decade. You have heard many things about his remarkable life, but I want to focus tonight on just one.

8   Aaron was part of an army of citizens that believes democracy only works when the citizenry are informed, when we know about our rights—and our obligations. An army that believes we must make justice and knowledge available to all—not just the well born or those that have grabbed the reigns of power—so that we may govern ourselves more wisely.

9   He was part of an army of citizens that rejects kings and generals and believes in rough consensus and running code.

10   We worked together on a dozen government databases. When we worked on something, the decisions weren't rash. Our work often took months, sometimes years, sometimes a decade, and Aaron Swartz did not get his proper serving of decades.

11   We looked at and poked at the U.S. Copyright database for a long time, a system so old it was still running WAIS. The government had—believe it or not—asserted copyright on the copyright database. How you copyright a database that is specifically called out in the U.S. Constitution is beyond me, but we knew we were playing with fire by violating their terms of use, so we were careful.

12   We grabbed that data and it was used to feed the Open Library here at the Internet Archive and it was used to feed Google Books. And, we got a letter from the Copyright Office waiving copyright on that copyright database. But before we did that, we had to talk to many lawyers and worry about the government hauling us in for malicious premeditated bulk downloading.

13   These were not random acts of aggression. We worked on databases to make them better, to make our democracy work better, to help our government. We were not criminals.

14   When we brought in 20 million pages of U.S. District Court documents from behind their 8 cent-per-page PACER pay wall, we found these public filings infested with privacy violations: names of minor children, names of informants, medical records, mental health records, financial records, tens of thousands of social security numbers.

15   We were whistle blowers and we sent our results to the Chief Judges of 31 District Courts and those judges were shocked and dismayed and they redacted those documents and they yelled at the lawyers that filed them and the Judicial Conference changed their privacy rules.

16   But you know what the bureaucrats who ran the Administrative Office of the United States Courts did? To them, we weren't citizens that made public data better, we were thieves that took $1.6 million of their property.

17   So they called the FBI, they said they were hacked by criminals, an organized gang that was imperiling their $120 million per year revenue stream selling public government documents.

18   The FBI sat outside Aaron's house. They called him up and tried to sucker him into meeting them without his lawyer. The FBI sat two armed agents down in an interrogation room with me to get to the bottom of this alleged conspiracy.

19   But we weren't criminals, we were only citizens.

20   We did nothing wrong. They found nothing wrong. We did our duty as citizens and the government investigation had nothing to show for it but a waste of a whole lot of time and money.

21   If you want a chilling effect, sit somebody down with a couple overreaching federal agents for a while and see how quickly their blood runs cold.

22   There are people who face danger every day to protect us—police officers and firefighters and emergency workers—and I am grateful and amazed by what they do. But the work that people like Aaron and I did, slinging DVDs and running shell scripts on public materials, should not be a dangerous profession.

23   We weren't criminals, but there were crimes committed, crimes against the very idea of justice.

24   When the U.S. Attorney told Aaron he had to plead guilty to 13 felonies for attempting to propagate knowledge before she'd even consider a deal, that was an abuse of power, a misuse of the criminal justice system, a crime against justice.

25   And that U.S. Attorney does not act alone. She is part of a posse intent on protecting property not people. All over the United States, those without access to means don't have access to justice and face these abuses of power every day.

26   It was a crime against learning when a nonprofit corporation like JSTOR, charged with advancing knowledge, turned a download that caused no harm and no damage into a $92 million federal case.

27   And the JSTOR corporate monopoly on knowledge is not alone. All over the United States, corporations have staked their fences on the fields of education: for-profit colleges that steal from our veterans, nonprofit standards bodies that ration public safety codes while paying million dollar salaries, and multinational conglomerates that measure the worth of scientific papers and legal materials by their gross margins.

28   In the JSTOR case, was the overly aggressive posture of the Department of Justice prosecutors and law enforcement officials revenge because they were embarrassed that—in their view at least—we somehow got away with something in the PACER incident? Was the merciless JSTOR prosecution the revenge of embarrassed bureaucrats because they looked stupid in the New York Times, because the U.S. Senate called them on the carpet?

29   We will probably never know the answer to that question, but it sure looks like they destroyed a young man's life in a petty abuse of power. This was not a criminal matter, Aaron was not a criminal.

30   If you think you own something and I think that thing is public, I'm more than happy to meet you in a court of law and—if you're right—I'll take my lumps if I've wronged you. But when we turn armed agents of the law on citizens trying to increase access to knowledge, we have broken the rule of law, we have desecrated the temple of justice.

31   Aaron Swartz was not a criminal, he was a citizen, and he was a brave soldier in a war which continues today, a war in which corrupt and venal profiteers try to steal and

hoard and starve our public domain for their own private gain.

32   When people try to restrict access to the law, or they try to collect tolls on the road to knowledge, or deny education to those without means, those people are the ones who should face the stern gaze of an outraged public prosecutor.

33   What the Department of Justice put Aaron through for trying to make our world better is the same thing they can put you through. Our army isn't one lone wolf, it is thousands of citizens—many of you in this room—who are fighting for justice and knowledge.

34   I say we are an army, and I use the word with cause because we face people who want to imprison us for downloading a database to take a closer look, we face people who believe they can tell us what we can read and what we can say.

35   But when I see our army, I see an army that creates instead of destroys. I see the army of Mahatma Gandhi walking peacefully to the sea to make salt for the people. I see the army of Martin Luther King walking peacefully but with determination to Washington to demand their rights because change does not roll in on the wheels of inevitability, it comes through continuous struggle.

36   When I see our army, I see an army that creates new opportunities for the poor, an army that makes our society more just and more fair, an army that makes knowledge universal.

Aaron's Army

37 When I see our army, I see the people who have created the Wikipedia and the Internet Archive, people who coded GNU and Apache and BIND and LINUX. I see the people who made the EFF and the Creative Commons. I see the people who created our Internet as a gift to the world.

38 When I see our army, I see Aaron Swartz and my heart is broken. We have truly lost one of our better angels.

39 I wish we could change the past, but we cannot. But, we can change the future, and we must.

40 We must do so for Aaron, we must do so for ourselves, we must do so to make our world a better place, a more humane place, a place where justice works and access to knowledge is a human right.

ON CRIME AND ACCESS TO KNOWLEDGE
UNPUBLISHED ESSAY
MARCH 30, 2013, SEBASTOPOL

Author's Note: In early February 2013, a British literary magazine asked me for 2-3 paragraphs about Aaron Swartz. I told them I couldn't really do justice to the issues in a piece that short, and after a month of back-and-forth turned in the attached 6,000-word essay. The editor declined, suggesting the piece might be more appropriate for the LA Times, which is how British literary editors nicely tell you not to quit your day job.

1   When my friend Aaron Swartz killed himself on January 11, 2013, it had been two years since he was arrested for downloading too many journal articles. Carmen Ortiz, the U.S. Attorney, had charged him with 13 felony counts, stating that "stealing is stealing whether you use a computer command or a crowbar."

2   Aaron had accessed the JSTOR database, a collection of over 1,800 academic journals that have been scanned and are available to subscribers, such as major research universities. For those that are allowed to access JSTOR, the service is a tremendous advance, allowing researchers to find, download, and read journal articles quickly and conveniently.

3   For those not fortunate enough to be at institutions such as Harvard or Oxford, JSTOR makes articles available at an average price of $21 per article, effectively locking the

rest of the world out from what science historian Lisbet Rausing called "the foundations of sociology, anthropology, geography, history, philosophy, classics, Oriental studies, theology, musicology, and the history of science." Aaron was deeply disturbed by JSTOR and other databases that erected walls between people and knowledge when he wrote in 2008 that "the world's entire scientific and cultural heritage, published over centuries in books and journals, is increasingly being digitized and locked up by a handful of private corporations."

4   Aaron didn't break into JSTOR, he used a valid JSTOR guest account available on the MIT campus, which runs an open network. Had he downloaded 1 article every day for 4.8 million days, there would have been no problem. Had he downloaded 100 articles every day for 48,000 days, that would have been fine as well, nobody would have noticed. But he downloaded 4.8 million articles in 100 days. Somewhere between 100 articles a day and 48,000 articles a day, Aaron crossed an invisible line.

5   He didn't release any of this information, it just accumulated on a disk drive, but the pace of the download brought an investigation by JSTOR, who called in the MIT network staff, who called in the police, and on January 6, 2011, Aaron was arrested. He was indicted on 2 counts of wire fraud, 4 counts of fraud and related activity in connection with computers, 5 counts of unlawfully obtaining information from a protected computer, and 1 count of recklessly damaging a protected computer. The charges had a maximum penalty of 35 years in prison and a $1 million fine.

6   This was a big deal, the kind of charges one brings against a gang stealing millions of credit cards. This wasn't a crowbar he was accused of using, he was accused of major capital crimes, an imminent danger to public order and public safety.

7   It is important to note that since Aaron's arrest, JSTOR has taken many steps to liberalize access to the archive, steps limited in large part by their inability to force reform with the publishers of these journals who have grown accustomed to immensely lucrative gross margins arising from their historical position as the designated intermediaries for the academic world and who set the per-article prices that JSTOR charges. One must also remember that JSTOR is a messenger, an intermediary, and if there is a fault here, that fault is ultimately the fault of the scholars who wrote those articles and allowed them to be locked up. It was a corruption of scholarship when the academy handed over copyright to knowledge so that it could be rationed in order to extract rents.

8   It is also important to note that much of the blame for the escalation of this situation rests with MIT, which is now conducting a broad internal investigation. The culture of experimentation and "hacking" at MIT stretches back to the earliest days of computers, and many students and faculty (including former visiting professors such as myself) are aghast at how they handled this situation. MIT clearly screwed this up.

9   No matter how one looks at this situation, the actions of JSTOR and MIT were the catalyst in a chain of events, and it was the actions of MIT and JSTOR that brought in the

federal prosecutors. Once the federal authorities were involved, even though JSTOR declined to press charges, there was no going back, and this led to a merciless 2-year prosecution to make an example of Aaron Swartz.

[…]

10 With Aaron's death, there has been an outpouring of emotion and analysis, an outpouring that has been a bit bewildering to many of us who knew him. Articles have appeared in mainstream media, congressional investigations have been promised and legislation has been introduced, countless blogs have analyzed the situation, memorial services have been held throughout the United States.

11 I was particularly struck by an essay published by the musician David Byrne, a man not known for worrying about the government. Byrne was very sympathetic to the idea that this knowledge should be more broadly available, and had clearly done a great deal of research to learn the facts. He placed Aaron's work in a long tradition of civil disobedience, and concluded by saying that when you break the law, you should expect to bear the consequences, and that Aaron had clearly not been prepared to face those consequences.

12 The U.S. Attorney, Carmen Ortiz, stated that she had not been trying to put Aaron away for the full 35 years possible for 13 felonies, she said that if Aaron pled guilty to all 13 charges, she would have asked for only 6 months in jail. Gandhi and Nehru spent years in jail, Martin Luther

King was arrested 29 times. If Aaron committed a crime, why not simply face the time?

13   Imagine if you will that you are accused of playing too loud in a bar, a crime David Byrne might have committed at some point. The police are called and you find yourself charged with 13 felonies for having endangered the public safety. The prosecutors say if you plead guilty to all 13 felonies, you will face only 6 months in jail. The federal courts also impose what is known as "supervised release," so after jail you will face 10 years of monthly drug tests and, because of the special nature of your crime, will not be allowed to touch a guitar or amplifier.

14   The shock that Aaron felt was the shock of disproportionality. If reading too many journal articles over the Internet is wrong, it is wrong like smoking a joint on campus or dismantling the Dean's car and rebuilding it in his office. It is the kind of act that, if it is indeed wrong, might perhaps result in a stern talking to, or a stint of community service. Instead, he faced accusations of computer and wire fraud, damages in the tens of millions of dollars.

15   A civil discussion about the role of public data was turned into a vicious criminal confrontation that looked like it would not only have put him in jail but and would also likely have resulted in a multiyear ban on using the Internet. Additional penalties imposed on felons include not voting, and somebody with 13 felonies isn't going to get a staff job in the White House or in Congress, the kinds of jobs Aaron aspired to. The disproportionality shocked Aaron and it shocked many others.

16   When you download a journal article, you are
presumed to be reading the article for purposes of
research. Aaron's alleged crime was not that he was
reading too many journal articles, the problem was that
Aaron was reading them with the help of a computer, that
he was reading them too fast. The crime he committed was
violating the vague yet complex "terms of use" that
governed access to the JSTOR database. The terms are an
implied agreement between the reader and the library
and they spell out the conditions under which one can
access this knowledge. The terms are adorned with a
great deal of vague yet fierce legal verbiage, such as "you
agree that you will not ... undertake any activity such as
computer programs that automatically download content."
The basic message is "behave yourself" but the terms
take over 9,000 words to express that message.

17   Technology changes rapidly, and the overlap between
a valid research technique and misbehavior is subject to
much debate. In the case of JSTOR, however, misbehavior
was defined like pornography, something that the
authorities would recognize when they saw it. What does it
mean to read an article in this day and age? Can one just
skim the section heads or just read the conclusion? Must
one comprehend what has been read? Can a computer
read on your behalf to identify salient facts? Proper use of
the articles that Aaron was reading was subject to his
agreement to the terms of use. All he had to do was
behave.

[…]

18   Aaron hadn't actually released any of those JSTOR journal articles when he was arrested. I'm convinced that Aaron had not made a decision to release those articles, and I am certain he would not have released them without a great deal of post-download analysis.

19   The implication was clear in the charges brought against him that he was about to let those journal articles loose in the wild, causing huge monetary damages. The prosecutor was convinced that the only reason that one would download journal articles was to redistribute them, and the indictment specifically charged that "Swartz intended to distribute these articles through one or more file-sharing sites."

20   Violating the terms of use, misbehaving by reading too many articles too quickly, is what turned Aaron's actions into a 13-felony federal crime. JSTOR called MIT who called the police. Then, even though JSTOR dropped all charges and asked the prosecutors not to pursue the case, the legal train had left the station and the prosecutors began a merciless 2-year pursuit. There was no posted speed limit on how many articles per day constituted an acceptable speed for downloading or reading. If, after the fact, the speed seemed inappropriate to prosecutors, they could and did invoke the full fury of the law. They called in those forces because an example had to be made of him, given his pattern of behavior in the past.

21   You see, he had done this before.

22   The last time Aaron had downloaded large numbers of journal articles was in 2008, when he downloaded 441,170 law review articles from Westlaw, a legal search service. He was trying to expose the practice of corporations such as Exxon funding a practice known as "for-litigation research," which consisted of lucrative stipends given to law professors who in turn produced articles penned specifically so they could be cited in ongoing litigation. In the case of Exxon, they were trying to reduce their $5 billion in punitive damages from the Exxon Valdez Oil Spill. Aaron didn't release any of the articles he downloaded, but the research he did was published in 2010 in a seminal article in the Stanford Law Review that exposed these ethically questionable practices in the legal academy.

23   Stringent terms of use on databases that may or may not have a public aspect to them are standard fare on the Internet. JSTOR is not the only walled garden with notices posted prominently on the gates. When Aaron and I downloaded the records of the U.S. Copyright Office database, a government-run registry of all works that have been filed for copyright protection, there were stringent terms of use prohibiting misuse and the threat of criminal prosecution was plainly stated on the web site. We ignored those dire notices since this was clearly, in our view, public data. Indeed, after the download had been completed, we got the Register of Copyrights to send us a letter confirming that the records were in the public domain and our actions had been perfectly appropriate.

24   In 2008, when we downloaded 20 million pages of public court records for U.S. District Courts from a government system known as PACER (ironically, for "Public Access to Electronic Court Records"), the terms of use were also exceedingly vague on what was allowed, but very specific about the criminal penalties. The login page for the court system warns in no uncertain terms:

25   *NOTICE: This is a restricted government web site for official PACER use only. Unauthorized entry is prohibited and subject to prosecution under Title 18 of the U.S. Code. All activities and access attempts are logged.*

26   Although all activities are logged, apparently the technical staff did not monitor those logs because the download of documents proceeded without a hitch for 2 months. We figured if there was a problem, somebody would say something and we'd stop. Instead, when the download was discovered, the Administrative Office of the U.S. Courts sent the FBI in to find us. The records were in fact public and we had done nothing wrong, a fact I explained to the two armed agents in the interrogation room.

27   The FBI had been called in by court administrators because they were charging 8 cents per page for access to the court database, a scheme that was bringing in over $100 million per year. Though there was no copyright on the data, from their point of view we had "taken" over $1.6 million of "their" documents without authorization because we had used a loophole to do our download. The courts had begun providing limited public access at 17

libraries around the country and we used that free access to pull in the documents and do a privacy audit, then after the documents had been cleaned up, to provide broader public access to public documents.

28   Sending the FBI out to try and find us was used instead of a more civilized dialogue, such as having the court administrators call one of us up to chat or perhaps holding a meeting on the subject. The fact that we were the ones that had their data was not a mystery, as we had sent registered letters to 32 Chief Judges, with copies to court administrators, containing details of massive privacy violations that we had uncovered.

29   Lawyers filing public documents with the court are under obligations to redact sensitive information, but these lawyers had left in tens of thousands of social security numbers, medical records, names and home addresses of minors, names of confidential informants, and other illegal information. The judges and their governing body, the Judicial Conference of the United States, had thanked us for our efforts and notified the U.S. Senate that they were changing their privacy procedures.

30   The resort to armed agents didn't seem like an attempt to try and understand how to serve the public better through wider distribution of public information or by determining how to make sure there were fewer privacy violations, it seemed like an attempt to intimidate, to protect a revenue stream, to prevent future access and future audits.

31   Signs that say "behave yourself" are standard fare on these databases and in many of the cases the terms are

clearly nonsense. If your vocation or avocation is trying to make public information public, you find yourself forced to evaluate the nature of the data and the nature of the prohibition and make a moral and legal choice as to whether it is right or safe to download information and then a second choice as to whether it is right or safe to release.

[…]

32 Why download the data at all? Why not simply send a letter to the officials and make the case that the information should be public? Because you can't tell until you've seen the information. In the case of JSTOR, one would have to ask how many of those articles were out of copyright because they were authored before 1923? How many of those articles were funded by federal research money requiring the release of any articles? Did the authors of those academic journal articles actually give a valid copyright assignment to the publishers or did the publishers perhaps presumptively assert ownership of electronic rights? Were some of the authors federal government employees who were required to release their work without restriction? You can't tell unless you look.

33 Aaron Swartz was accused of abusing his library card. He was charged with the intent to propagate knowledge. Prosecutors were convinced he would upload his trove of 4.8 million documents to "file-sharing networks." Would he have released that database? I think that if he had

looked at the data and could have convinced himself and others that there was a sound moral and legal argument for release, he would have done so in a minute, but he was far from making that decision.

34   Instead of releasing the data, he might have used his analysis to approach JSTOR either privately or publicly, and argue his case to them armed with real data. Perhaps he would have approached the students, alumni, or faculty of universities such as Harvard to try and get them to press his case. You can't tell what is in a database unless you look, and I know in the past we had frequently looked at data and concluded we didn't have a sound basis for release and we held our fire.

35   Databases such as JSTOR, and certainly the government databases such as court or copyright records, all have a strong public component. In many cases, this is knowledge that is supposed to be public, that is meant to save lives and educate our children, or ensure equal protection under the law. Most people will recognize that some knowledge belongs to the public, no matter how much they support strong copyright protection for creative works. The question is where to draw the line between public and private.

36   There is indeed private knowledge where the decision to release is left up to the rights holder. In 1909, when Mohandas K. Gandhi published Hind Swaraj ("Indian Home Rule"), he could have registered the work for copyright protection, but choose instead to print "No Rights Reserved" on the cover. If you buy a copy of the Centenary Edition of Hind Swaraj from Cambridge

University Press, you'll note that they carefully skirt the issue of any copyright on the text authored by Gandhi, who felt that "writings in the journals which I have the privilege of editing must be the common man's property."

37   Gandhi's waiver of certain rights is very similar to a Creative Commons license, a system which Aaron helped develop that allows authors to waive some or all of their rights. A license might be granted, for example, that allows use of a work but demands attribution, or prohibits derivative works or commercial use. "No Rights Reserved" is the same as the license known as "CC-Zero," which allows a publisher to waive all rights on a work. My organization has used the CC-Zero license on thousands of government videos and millions of government documents to make clear that we reserve no rights in this information. Though entitled to copyright, many authors (including many if not most of the authors of academic journal articles) would choose the pursuit of knowledge over restrictions on use. For private knowledge, the decision is left to the author or publisher.

38   Certain kinds of information, however, are owned by everybody not by somebody. These works are not entitled to copyright protection, or if they have copyright owned by the state (known as "Crown Copyright" in the United Kingdom), there is an implied license that allows people to read and speak the texts. Examples of such documents that are owned by the people include edicts of government, such as laws and court opinions. JSTOR was clearly a hybrid, with some copyright material, but also a large public component that certainly should have been

available to a broader audience on moral grounds, and Aaron obviously felt it was worth a look to see if there were also legal grounds that might help prompt the release of these reservoirs of knowledge.

39   After Aaron was arrested, his lawyers told him not to talk to his friends because they could be required to testify against him. He had not stated publicly what he was going to do with all those journal articles, and he never told me. I only had one exchange with him on the subject of JSTOR. A few days after he was arrested, he sent me a note and asked if I knew Kevin Guthrie, who was the CEO of JSTOR. I didn't place the name right away, so I sent him back a flip answer, saying "No, but I'm a big fan of Arlo and Woody." I then looked up Guthrie's name, made the connection to JSTOR, and sent him back a long note with my theories about academic journal articles. He responded with a one-liner: "I've been pursuing research along similar lines."

40   I didn't know then he had just been arrested. I only found out a few months later when the New York Times called, told me what had happened, and asked if I had any comment. Aaron and I didn't talk a lot after that because we were both warned that I was as likely a person to be required to testify as anybody. (I was never called.) I followed his case on the court system, kept up with his Twitter feed, saw him occasionally logon to Skype, but we didn't talk, and I've very much regretted that.

41   Aaron did talk to one friend of ours after his arrest, a prominent member of the technical world who was one of the people responsible for revolutionizing how

presidential campaigns are conducted in the U.S., first with the Dean campaign, then for Obama's 2008 run. Aaron had a long talk with him and told him his plans to repeat his Westlaw analysis, this time using the JSTOR database to examine whether climate change research had been influenced by corporate money.

42   Does this mean that he wouldn't have released the documents into the wild? No, he might have, but I am convinced he had not made that decision. In our world, a release decision is something you make after you've built a thing, not a wish or a dream. You hope what you've built is good enough, but if you play in the big leagues, sometimes you build things that aren't ready, and those don't get released.

43   It was very clear, however, that Aaron wanted to analyze those journal articles, a use that was clearly beneficial. If Aaron was going to release the data, his prior pattern of behavior had been to go find somebody like me to determine how to release it properly with a strong legal argument to back the release. The fact that he had 4.8 million articles on his laptop and had not talked to anybody yet argues strongly against any intent to distribute. Even if there was an intent, he had not distributed the data. No damage had been done, certainly not a capital crime worthy of 13 felonies.

[…]

44   A database very similar to JSTOR is one that I've been working with for the past year. Building codes, fire codes,

worker protection standards, and numerous other technical public safety documents are created under copyright and then mandated by government. The cost of these edicts of government is breathtaking, with prices of hundreds of pounds per document not at all uncommon. Because our technologically sophisticated world is so complex, you need to read hundreds of such laws for even the narrowest of niches, such as the safety of implantable medical device or protocols for testing water or ecologically sound (and legally mandated) practices for packaging goods. Not all these technical standards become law, but many of them do, and these documents regulate everything from the safety of factory and agriculture machinery to the transport of hazardous goods to standards for safe construction of schools.

45   An example of such a public safety standard I released just before Aaron's death is the European building code, known as the Eurocode, which is required to be enacted in full without change by every European Union country and replaces the previous system of national building codes. While the official Eurocode site has a few presentations and reports, the code is only made available through national standards bodies. If you are in the United Kingdom, this standard costs £11,674 for the 58 parts, payable to the nonprofit crown entity known as the British Standards Institution, plus £3061 for the UK national annexes, which detail any local additions to the code.

46   The Eurocode spells out the requirements of construction for residential, commercial, and industrial buildings, covering topics such as fire safety, structural

integrity of concrete structures, seismic safety, integrity under snow loads and wind actions, traffic loads on bridges, safety of silos and tanks, and the proper construction of pipelines, towers, masts, and chimneys. Every home owner and building manager is presumed to know and obey this building code. This is not a document just for specialists, it is a law of general applicability, and ignorance of the law is no excuse.

47   For republishing the Eurocode and 469 other European standards on our web site without permission, I know that I might find myself in the same position as Aaron. Because our site includes 10,062 standards from 24 countries and 6 international organizations, standards we purchased for $180,410, I know to many this looks just like JSTOR. Unlike JSTOR, however, we've released those documents, and the reason we released them is we believe we have a right to do so.

48   Public promulgation of the law is a fundamental aspect of a doctrine known as the rule of law, a doctrine reflected in constitutions and treaties throughout the world. The rule of law states that we will govern ourselves by rules that are set out in advance, known by all, and enforced in open and fair courts that apply the same law to all. The Eurocode is a law, and the idea that the documents with the force of law should be freely available is anchored in basic human rights documents such as the European Convention for the Protection of Human Rights, which begins with a preamble affirming that all Europeans "have a common heritage of political traditions, ideals, freedom

and the rule of law" and goes on to state this includes "no punishment without law."

49 The rule of law is not a recent invention, you can see it clearly stated in Magna Carta when it affirms "To no one will we sell, to no one will we deny or delay, right or justice." This provision of Magna Carta is not a historical artifact, it is still binding law in the United Kingdom and in countries such as New Zealand and the United States. If it costs £11,674 to read an important law, have we not put a price on right and justice?

50 As dangerous as high prices to read the law are, prohibitions against copying the documents and sharing them with other citizens are more insidious. The European Convention plainly addresses this in Article 10, which states: "Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers."

51 Again, one might ask why take the risk to republish a £11,674 document such as the Eurocode, why not simply ask the authorities politely? Quite simply, this is not a conversation the pseudo-government entities that publish these documents will entertain. There are many billions of dollars per year in revenue at stake, there are many million-dollar salaries. The idea that the sums being extracted from citizens for the privilege of reading the law is illegal is not a conversation they wish to have.

52 Though we've written numerous letters to government officials asking to discuss the matter, even officials known

for their progressive positions on transparency have not seen fit to answer, or have declined to take up the matter. In the United Kingdom, for example, we wrote to the Rt. Hon. Francis Maude, who as Minister for the Cabinet Office has spearheaded a very impressive digital transformation of his government, but his staff responded that "we are primarily concerned with the publication of open data" and suggested we take the matter up with the Ministry of Justice or the National Archives. The Ministry of Justice is of course charged with representing the government, and is thus more likely to want to hear from crown entities such as the British Standards Institution. The National Archives runs one of the world's most impressive legislative sites, but they can only archive what they are handed. The question of what laws should be public is not one for a specialized agency, but rather one that must be discussed within government as a whole, such as the Parliament or the Cabinet.

53   There is only one way such a document becomes available, and that is when people make it available and begin the conversation as a matter of reality not theory. Access has to start somewhere, and we published these documents not from some ideological position but because we believe sincerely it is our legal right to do so. The decision was not rash, it was based in an intense study of the rule of law, human rights, and the essential role in public safety of each document we published.

[…]

54 Reading the law, or reading scholarly literature, is a fundamental human right. Being able to speak the law, or repeat the journal articles, is just as important. It is by imparting knowledge that we make it grow. When we publish a legal document, we do much more than simply scan in a piece of paper. We retype and reset the standards to make them more accessible on the web. We redraw the diagrams and code the formulas. We optimize the documents so search engines can find them. We do what we wish the governmental bodies would do to provide public access, and we do so in the hope that they will provide that access themselves.

55 That our publication of a document such as the Eurocode should be the subject of civil conversation, even the possibility of a suit in a civil court is well recognized. Aaron recognized that if he were to decide to publish pieces of JSTOR, that would also perhaps be the subject of a conversation in a civil court. That efforts to discover the public nature of a database such as JSTOR became a criminal matter was a shock. That publication of edicts of government should perhaps become a criminal matter is a shock.

56 Public data matters not because of an ideology that information must be free or that copyright is evil. Public data matters because making certain kinds of information available for all to know (and in some cases obey) makes our society work properly. During the American revolution, there was great fear that the radical experiment in democracy that was being proposed would

be a failure since the people could not be trusted to govern themselves.

57   To rebut this canard, John Adams sent in four columns to the Boston Gazette under the title "A Dissertation on the Canon and the Feudal Law," in which he argued that the if the citizenry were educated and well-informed, democracy would not only work, it would be a far better form of government than those run by bishops or barons. Adams said that if we believe "truth, liberty, justice, and benevolence are the everlasting basis of law and government," then we must arm our citizens with knowledge. Public information shouldn't be a conceal-carry privilege for the rich. Adams said that for democracy to work, we must:

58   *"Let the public disputations become researches into the grounds and nature and ends of government, and the means of preserving the good and demolishing the evil. Let the dialogues, and all the exercises, become the instruments of impressing on the tender mind, and of spreading and distributing far and wide, the ideas of right and the sensations of freedom.*

59   *In a word, let every sluice of knowledge be opened and set a-flowing."*

[…]

60   Aaron Swartz was arrested for reading too many articles with an intent to propagate that knowledge, of threatening to open a sluice that once open could never be closed.

The commotion and anger that resulted from his death was not just because a brave man had been beaten down, it is because we see that we could all face that same fury for the abuse of knowledge that might be, should be, or is public. That this would be a crime shatters the rule of law.

What astounded me when news of his death became public was just how many people stood up and told stories of their close collaborations with him. All over the world, one after another people started speaking up and saying how he had helped them in their work, had touched their lives. He had a reputation as a loner, as many of us who spend our lives behind a computer do, but he was what we call on the net "wired," a person who seemed to know everybody.

The Internet is a big place. But, for those of us who have spent our lives helping to build it, the Internet is still a small world. Aaron was a young prodigy, he made his mark at the age of 14 by creating a basic Internet protocol known as RSS, a mechanism for one web site to notify a browser or other web site that new content was available. When he was 16, he was working with Lawrence Lessig to help create Creative Commons, open licenses that have enabled content to be made freely and easily available on sites such as YouTube, Wikipedia, and Flickr. He worked with Brewster Kahle and the Internet Archive to create their Open Library system, a repository of 2.5 million books. He worked with the best web designers in the business, he helped build tools for a new generation of political activists. He worked with Tim Berners-Lee on the creation of a new generation of web protocols known as

the Semantic Web. When he died at 26, I had been working with him for 10 years and I rarely noticed his age, he was simply a colleague and he was as often our teacher as we were his.

63   Aaron did many things in his short life, and I had nothing to do with his best work, like the day he got the entire Internet to go dark in protest over an Internet censorship bill. I did work with him, however, on the kinds of projects, like JSTOR and the Court database, that ended up being the most terrifying.

64   It is tempting to lay Aaron's work in with Wikileaks or the Anonymous collective, but those efforts want to change or expose or destroy the system. Aaron wanted to work within the system, to hack it, to make it better, to make it do things that that had not been done but were possible and desirable. If you disagreed with him, he wanted to talk about it, to convince you he was right. He didn't want to destroy JSTOR, he wanted to make JSTOR better, to realize the full potential of JSTOR.

65   When Aaron killed himself, Tim Berners-Lee spoke for many when he posted a message that read: "Aaron dead. World wanderers, we have lost a wise elder. Hackers for right, we are one down. Parents all, we have lost a child. Let us weep." I wept.