UNITED STATES DISTRICT COURT
DISTRICT OF MASSACHUSETTS

UNITED STATES OF AMERICA	**Crim. No.** 11-cr-10260
v.	**VIOLATIONS:**
AARON SWARTZ,	**18 U.S.C. § 1343 (Wire Fraud)**
Defendant	**18 U.S.C. § 1030(a)(4) (Computer Fraud)**
	18 U.S.C. § 1030(a)(2), (c)(2)(B)(iii) (Unlawfully Obtaining Information from a Protected Computer)
	18 U.S.C. § 1030(a)(5)(B), (c)(4)(A)(i)(I),(VI) (Recklessly Damaging a Protected Computer)
	18 U.S.C. § 2 (Aiding and Abetting)
	18 U.S.C. § 981(a)(1)(C), 28 U.S.C. § 2461(c), and 18 U.S.C. §982(a)(2)(B) (Criminal Forfeiture)

INDICTMENT

The Grand Jury charges that at all relevant times:

PARTIES

1. The Massachusetts Institute of Technology ("MIT") was and continued to be a leading research and teaching university located in Cambridge, Massachusetts.

2. JSTOR, founded in 1995, was and continued to be a United States-based, not-for-profit organization that provides an online system for archiving and providing access to academic journals. It provides searchable digitized copies of over 1,000 academic journals, dating back for lengthy periods of time.

3. JSTOR's service is important to research institutions and universities because it can be extraordinarily expensive, in terms of both cost and space, for a research or university library to maintain a comprehensive collection of academic journals. By digitizing extensive, historical collections of journals, JSTOR enables libraries to outsource the journals' storage,

ensures their preservation, and enables authorized users to conduct full-text, cross-disciplinary searches of them. JSTOR has invested millions of dollars in obtaining and digitizing the journal articles that it makes available as part of its service.

4. JSTOR generally charges libraries, universities, and publishers a subscription fee for access to JSTOR's digitized journals. For a large research university, this annual subscription fee for JSTOR's various collections of content can cost more than $50,000. Portions of the subscription fees are shared with the journal publishers who hold the original copyrights. In addition, JSTOR makes some articles available for individual purchase. Publishers decide which articles can be purchased individually and set fees for their articles. JSTOR facilitates the purchase of these articles from the archive on behalf of the participating publishers.

5. JSTOR did not permit users:

a. to download or export content from its computer servers with automated computer programs such as web robots, spiders and scrapers;

b. to download all of the articles from any particular issue of a journal; or

c. to make other than personal use of individually downloaded articles.

6. JSTOR notified its users of these rules, and users accepted these rules when they chose to obtain and use JSTOR's content.

7. JSTOR provided MIT with its services and content for a fee.

8. MIT, in turn, made JSTOR's services and content available to its students, faculty, and employees. MIT also allowed guests of the Institute to have the same access as its students, faculty, and employees for short periods of time while they were on campus.

9. JSTOR's computers were located outside the Commonwealth of Massachusetts, and thus any communications between JSTOR's computers and MIT's computers in Massachusetts crossed state boundaries. JSTOR's computers were also used in and affected interstate and foreign commerce.

10. Aaron Swartz lived in the District of Massachusetts and was a fellow at Harvard

University's Center for Ethics. Although Harvard provided Swartz access to JSTOR's services and archive as needed for his research, Swartz used MIT's computer networks to steal well over 4,000,000 articles from JSTOR. Swartz was not affiliated with MIT as a student, faculty member, or employee or in any other manner other than his and MIT's common location in Cambridge. Nor was Swartz affiliated in any way with JSTOR.

OVERVIEW OF THE OFFENSES

11. Between September 24, 2010, and January 6, 2011, Swartz contrived to:

a. break into a restricted computer wiring closet at MIT;

b. access MIT's network without authorization from a switch within that closet;

c. connect to JSTOR's archive of digitized journal articles through MIT's computer network;

d. use this access to download a major portion of JSTOR's archive onto his computers and computer hard drives;

e. avoid MIT's and JSTOR's efforts to prevent this massive copying, measures which were directed at users generally and at Swartz's illicit conduct specifically; and

f. elude detection and identification;

all with the purpose of distributing a significant proportion of JSTOR's archive through one or more file-sharing sites.

MEANS OF COMMITTING THE OFFENSES

12. Swartz alone, or in knowing concert with others unknown to the grand jury, (hereafter simply "Swartz" in this section) committed these offenses through the means described below.

September 24 through 27, 2010

13. On September 24, 2010, Swartz purchased an Acer laptop computer from a local

computer store with the intent of using it to automatically and systematically harvest JSTOR's archive of digitized journal articles.

14. Later that day, Swartz connected the Acer computer to MIT's computer network from a location in Building 16 at MIT and registered under a pseudonym with MIT's computer network as a guest. MIT offers campus guests short-term service on its computer network. Campus guests must register on the MIT network and are limited to a total of fourteen days per year of network service.

15. Swartz registered on the network using identifiers chosen to hide his identity as the computer's owner and user.

a. The computer was registered under the fictitious guest name "Gary Host."

b. The computer's client name was specified as "ghost laptop." A computer's client name helps to identify it on a network and can be chosen by its user. In this case, the name was simply created by abridging the pseudonym "Gary Host," combining the first initial "g" with the last name "host."

c. The fictitious "Gary Host's" e-mail address was identified as "ghost@mailinator.com." This was a "throwaway" e-mail address. Mailinator is a free, disposable e-mail service that allows a user to create a new e-mail address as needed, without even registering the address with Mailinator. Mailinator provides this service for users to have an anonymous and temporary e-mail address. Mailinator accepts mail for any e-mail address directed to the mailinator.com domain without need for a prior registration, and it allows anyone in the world to read that mail without having to create an account or enter a password. All mail sent to mailinator.com is automatically deleted after several hours whether read or not.

16. On September 25, 2010, Swartz used the Acer laptop to systematically access and rapidly download an extraordinary volume of articles from JSTOR. He used a software program

to automate the downloading process so that a human being would not need to keep typing in the archive requests. The program was also designed to sidestep or confuse JSTOR's efforts to prevent this behavior.

17. These rapid and massive downloads and download requests impaired computers used by JSTOR to service client research institutions and threatened to misappropriate its archive.

18. As JSTOR, and then MIT, became aware of these efforts to steal a vast proportion of JSTOR's archive, each took steps to block the flow of articles to Swartz's computer and thus to prevent him from redistributing them. Swartz, in turn, repeatedly altered the appearance of his Acer laptop and the apparent source of his automated demands to get around JSTOR's and MIT's blocks against his computer.

a. On the evening of September 25, 2010, JSTOR blocked the computer's access to its network by refusing communications from the computer's assigned IP address. An IP (short for "Internet Protocol") address is a unique numeric address used by a computer on the Internet. Every computer attached to the Internet must be assigned an IP address so the Internet traffic sent from and directed to that computer can be directed properly from the source to its destination. Most Internet service providers control a range of IP Addresses. MIT controls all IP addresses that begin with the number 18. In this case, the computer had been assigned an IP address of 18.55.6.215, and JSTOR blocked communications from that IP address.

b. On September 26, 2010, Swartz obtained for his computer a new IP address on the MIT network – 18.55.6.216 – and began again to download an extraordinary volume of articles from JSTOR. Accesses from this address continued until the middle of the day, when JSTOR spotted and blocked this IP address as well. Because the exploits on September 25 and 26 were both

launched from MIT IP addresses beginning with 18.55.6 , and because computers used by JSTOR to service client research institutions were again impaired and its archive at risk of misappropriation, on September 26, 2010, JSTOR began blocking a much broader range of IP addresses. As a result, legitimate JSTOR users at MIT were denied access to JSTOR's archive until September 29, 2010.

c. Notified by JSTOR of what was happening, MIT sought to block Swartz more specifically. It did so by prohibiting the Acer laptop from being assigned an IP address on MIT's network. When a user plugs his computer into the wired network on MIT's campus, his computer's MAC address is used to determine whether he has been authorized to use the network. A MAC address is a unique identifier assigned to a computer network interface, in this case, the Acer laptop's network interface card. A MAC address most often is assigned by the manufacturer of the network interface card and therefore generally remains constant on the device. Although a MAC address is intended to be a permanent and globally unique identification, a user with the right knowledge can change the MAC address, an action referred to as "MAC address spoofing," as discussed below.

d. As part of the registration process, "Gary Host's" computer, i.e., the Acer laptop, had identified its network interface's MAC address as 00:23:5a:73:5f:fb. Consequently, on September 27, 2010, MIT deactivated the guest registration for the "ghost laptop" by barring any network interface with that MAC address from being assigned a new IP address.

19. MIT banned the Acer laptop from its network under and consistent with its own computer use rules, which required users to:

a. use the network to support MIT's research, education, and MIT administrative activities, or at least to not interfere with these activities;

b. maintain the system's security and conform to applicable laws, including copyright laws; and

c. conform with rules imposed by any networks to which users connected through MIT's system.

Guest users of the MIT network agreed to be bound by the same rules that applied to students, faculty, and employees. These rules explicitly notified users that violations could lead to state or federal prosecution.

October 2 through 9, 2010

20. Despite knowing that his computer had been blocked from JSTOR's and MIT's networks, Swartz sought and obtained another guest connection on MIT's network, again for his Acer laptop less than a week later, on October 2, 2010.

21. Once again, Swartz registered the Acer latop on the network using identifiers chosen to avoid identifying Swartz as the computer's owner and user:

a. The computer was once again registered under the fictitious name "Gary Host" and the client name "ghost laptop."

b. To evade the MAC address block, Swartz spoofed the computer's MAC address, manipulating it from 00:23:5a:73:5f:fb to 00:23:5a:73:5f:fc (the final "b" became a "c").

c. By re-registering the "ghost laptop," Swartz ensured that it was assigned a new IP address. By obtaining a new IP address, Swartz disassociated his rogue computer from the IP addresses used to exploit JSTOR in September.

22. On October 8, 2010, Swartz connected a second computer to MIT's network and registered as a guest, using similar naming conventions: the computer was registered under the name "Grace Host," the computer client name "ghost macbook," and the throw-away e-mail address "ghost42@mailinator.com."

23. The next day, October 9, 2010, Swartz used both the "ghost laptop" and the

"ghost macbook" to systematically and rapidly access and download an extraordinary volume of articles from JSTOR. The pace was so fast that it brought down some of JSTOR's computer servers.

24. In response, JSTOR blocked the entire MIT computer network's access to JSTOR for several days, beginning on or about October 9, 2010.

November and December, 2010

25. During November and December, 2010, Swartz used the "ghost laptop" (i.e., the Acer laptop) at MIT to make over two million downloads from JSTOR. This is more than one hundred times the number of downloads during the same period by all the legitimate MIT JSTOR users combined. Of the downloads, approximately half were research articles, with the remainder being reviews, news, editorials, and miscellaneous documents.

26. This time around, Swartz circumvented MIT's guest registration process altogether when he connected to MIT's computer network. By this point, Swartz was familiar with the IP addresses available to be assigned at the switch in the restricted network interface closet in the basement of MIT's Building 16. Swartz simply hard-wired into the network and assigned himself two IP addresses. He hid the Acer laptop and a succession of external storage drives under a box in the closet, so that they would not be obvious to anyone who might enter the closet.

January 4 through 6, 2011

27. On January 4, 2011, Aaron Swartz was observed entering the restricted basement network wiring closet to replace an external hard drive attached to his computer.

28. On January 6, 2011, Swartz returned to the wiring closet to remove his computer equipment. This time he attempted to evade identification at the entrance to the restricted area. As Swartz entered the wiring closet, he held his bicycle helmet like a mask to shield his face, looking through ventilation holes in the helmet. Swartz then removed his computer equipment from the closet, put it in his backpack, and left, again masking his face with the bicycle helmet

before peering through a crack in the double doors and cautiously stepping out.

29. Shortly thereafter, Swartz connected the Acer latop to MIT's network in a different building, again registering on the network using identifiers chosen to avoid identifying Swartz as the computer's owner and user.

a. The computer was registered under the fictitious name "Grace Host" and the client name "ghost laptop."

b. To evade the block on the computer's MAC address, Swartz had spoofed (manipulated) its MAC address a second time, changing it from the blocked 00:23:5a:73:5f:fb to 00:4c:e5:a0:c7:56.

c. By re-registering the "ghost laptop," Swartz ensured that it was assigned a new IP address. By obtaining a new IP address for his rogue computer, Swartz disassociated it from the IP addresses used to exploit JSTOR up to that point.

30. Swartz had a software program named "keepgrabbing.py" installed on the Acer laptop. Keepgrabbing.py was designed to download .pdf files from jstor.org and sidestep or confuse JSTOR's efforts to prevent the behavior.

31. When MIT Police spotted Swartz on the afternoon of January 6, 2011 and attempted to question him, he fled with a USB drive that contained the program "keepgrabbing2.py." "Keepgrabbing2.py" had distinct similarities to "keepgrabbing.py."

32. In all, Swartz stole approximately 4.8 million articles, a major portion of the total archive in which JSTOR had invested. Of these, approximately 1.7 million were made available by independent publishers for purchase through JSTOR's Publisher Sales Service.

33. Swartz intended to distribute a significant portion of JSTOR's archive of digitized journal articles through one or more file-sharing sites.

COUNT 1
Wire Fraud
18 U.S.C. §§ 1343 & 2

34. The Grand Jury realleges and incorporates by reference the allegations in paragraphs 1-33 of this Indictment and charges that:

From on or about September 24, 2010, through January 6, 2011, or thereabout, in the District of Massachusetts and elsewhere, the defendant,

AARON SWARTZ,

having devised and intended to devise a scheme and artifice to defraud and for obtaining property — namely, journal articles digitized and distributed by JSTOR, and copies thereof — by means of material false and fraudulent pretenses, representations, and promises, transmitted and caused to be transmitted by means of wire communication in interstate commerce writings, signs, signals, and pictures — namely, communications to and from JSTOR's computer servers — for the purpose of executing the scheme, and aided and abetted the same.

All in violation of Title 18, United States Code, Sections 1343 and 2.

COUNT 2
Computer Fraud
18 U.S.C. §§ 1030(a)(4) & 2

35. The Grand Jury realleges and incorporates by reference the allegations in paragraphs 1-33 of this Indictment and charges that:

From on or about September 24, 2010, through January 6, 2011, or thereabout, in the District of Massachusetts and elsewhere, the defendant,

AARON SWARTZ,

knowingly and with intent to defraud, accessed a protected computer — namely, a computer on MIT's network and a computer on JSTOR's network — without authorization and in excess of authorized access, and by means of such conduct furthered the intended fraud and obtained things of value — namely, digitized journal articles from JSTOR's archive — and aided and abetted the same.

All in violation of Title 18, United States Code, Sections 1030(a)(4) and 2.

COUNT 3
Unlawfully Obtaining Information from a Protected Computer
18 U.S.C. §§ 1030(a)(2), (c)(2)(B)(iii) & 2

36. The Grand Jury realleges and incorporates by reference the allegations in paragraphs 1-33 of this Indictment and charges that:

From on or about September 24, 2010, through January 6, 2011, or thereabout, in the District of Massachusetts and elsewhere, the defendant,

AARON SWARTZ,

intentionally accessed a computer — namely, a computer on MIT's computer network and a computer on JSTOR's network — without authorization and in excess of authorized access, and thereby obtained from a protected computer information whose value exceeded $5,000 — namely, digitized journal articles from JSTOR's archive — and aided and abetted the same.

All in violation of 18 U.S.C. §§ 1030(a)(2), (c)(2)(B)(iii) and 2.

COUNT 4
Recklessly Damaging a Protected Computer
18 U.S.C. §§ 1030(a)(5)(B), (c)(4)(A)(i)(I),(VI) & 2

37. The Grand Jury realleges and incorporates by reference the allegations in paragraphs 1-33 of this Indictment and charges that:

From on or about September 24, 2010, through January 6, 2011, or thereabout, in the District of Massachusetts and elsewhere, the defendant,

AARON SWARTZ,

intentionally accessed a protected computer — namely, a computer on MIT's computer network and a computer on JSTOR's network — without authorization, and as a result of such conduct recklessly caused damage to MIT and JSTOR, and, during a 1-year period, caused loss aggregating at least $5,000 in value and damage affecting at least 10 protected computers, and aided and abetted the same.

All in violation of Title 18, United States Code, Sections 1030(a)(5)(B), (c)(4)(A)(i)(I),(VI) & 2.

FORFEITURE ALLEGATIONS

(18 U.S.C. § 981(a)(1)(C), 28 U.S.C. § 2461(c), and 18 U.S.C. §982(a)(2)(B))

38. Upon conviction of the offense alleged in Count One of the Indictment, the defendant,

AARON SWARTZ,

shall forfeit to the United States, pursuant to 18 U.S.C. § 981(a)(1)(C) and 28 U.S.C. § 2461(c), any property, real or personal, that constitutes, or is derived from, proceeds traceable to the commission of the offense.

39. Upon conviction of one or more of the offenses alleged in Counts Two through Four of the Indictment, the defendant,

AARON SWARTZ,

shall forfeit to the United States, pursuant to 18 U.S.C. § 982(a)(2)(B) any property constituting, or derived rom, proceeds obtained directly or indirectly as a result of the commission of the offenses.

40. If any of the property described in paragraphs 38 and 39 hereof as being forfeitable pursuant to 18 U.S.C. § 981(a)(1)(C), 28 U.S.C. § 2461(c), and 18 U.S.C. § 982(a)(2)(B) as a result of any act or omission of the defendant --

a. cannot be located upon the exercise of due diligence;

b. has been transferred to, sold to, or deposited with a third party;

c. has been placed beyond the jurisdiction of this Court;

d. has been substantially diminished in value; or

e. has been commingled with other property which cannot be divided without difficulty;

it is the intention of the United States, pursuant to 21 U.S.C. § 853(p), as incorporated by 28 U.S.C. § 2461(c), to seek forfeiture of all other property of the defendant up to the value of the property described in paragraphs 38 and 39 above.

All pursuant to Title 18, United States Code, Sections 981(a)(1)(C) and 982(a)(2)(B), and Title 28, United States Code, Section 2461(c).

A TRUE BILL

David R Thompson
Foreperson of the Grand Jury

Assistant United States Attorney

Date: 7-14-11

DISTRICT OF MASSACHUSETTS

July 14, 2011

Returned into the District Court by the Grand Jurors and filed.

Deputy Clerk

2:09p