From: ██████████████████████ @ithaka.org>
Sent: Wednesday, May 11, 2011 9:36 PM (GMT)
To: Heymann, Stephen (USAMA) <Stephen.Heymann@usdoj.gov>
Subject: FW: Could you check one thing in the morning for me?

Hi Steve,

Please see ██████ response below. As you can see, this is fairly complex!

Let me know if you need to discuss further with ████.

████████

From: ████████
Sent: Wednesday, May 11, 2011 4:59 PM
To: ██████████████████████
Subject: RE: Could you check one thing in the morning for me?

Thanks ████████

The quick answer is no, apparently it is nowhere near that simple. Hopefully cogent points...

1. Backfile (non-CSP eBack Swap) typically look like this
   http://www.jstor.org/stable/pdfplus/40184736.pdf
2. Frontfile and CSP eBack Swap typically looks like this
   http://www.jstor.org/stable/pdfplus/10.1525/auk.2011.128.1.fm.pdf, but can sometimes look like
   this http://www.jstor.org/stable/pdfplus/10.1641/0006-3568%282003%29053%5B0915%
   3ATDOOF%5D2.0.CO%3B2.pdf (not a joke, surprisingly)

And, while scenario #1 sort of fits the description he provides, I don't think that is what is going on here.
I don't know of any .pdf extensions that end with an X.

We'll need Content Management to frame this discussion if we need to produce a definitive list of how
they are constructed and what they can and cannot look like.

Here's what I am thinking.

3. Could be red herring, could not be, but the question he asked immediately made me think of
   ISSNs
   a. They often contain an X and only at the end.
   b. Many also have leading zeros like 00023456, which get dropped automatically in Excel,
      etc (thus the 4-7 character part of his analysis).
   c. ISSNs are ALWAYS 8 numbers or 7 numbers followed by an X. Again, might look like
      less if leading zeros are no longer present
   d. When you save a PDF, it saves it only as these odd strings seen above that aren't very
      helpful in identifying what it is you actually saved (like, it doesn't save as Title-Author or
      something human readable)
   e. If, however, you wanted to save our files in a systematic way that would allow you to
      organize, index, etc., you might use the ISSN as a way to rename the files as you are
      downloading them (ISSNs are unique to journal titles and even change when split into
      two journals, etc., but are constant for a title while its run is uninterrupted, aka, a good

basis for file naming.)

Is it possible Stephen is looking at the files as they are stored somewhere and not at our logs or records?

That is my guess, he's seeing the naming convention used by the intruder to save and organize acquired files, not JSTOR's .pdf naming convention.

Best,

███████████████

JSTOR | Portico

███████@ithaka.org
██████

From: Heymann, Stephen (USAMA) [mailto:Stephen.Heymann@usdoj.gov]
Sent: Tuesday, May 10, 2011 6:18 PM
To: ████████
Subject: Could you check one thing in the morning for me?

Am I correct that the naming convention for the .pdf files of your digitized journal articles is to use either a 5-8 digit number or a 4-7 digit number followed by the letter "x"?

Steve