

To: Jeffrey Epstein[jeevacation@gmail.com]
From: Lesley Groff
Sent: Tue 6/22/2010 6:59:35 PM
Subject: Victoria Stodden

Title: Re: Jeffrey Epstein

From: Victoria Stodden [mailto: [REDACTED]]
Sent: Tuesday, June 22, 2010 2:58 PM
To: Lesley Groff; Jeffrey Epstein
Cc: [REDACTED]
Subject: Re: Jeffrey Epstein

Jeffrey,

* This is a fun summer for 1) *travel and talks*:

June 15-18: Workshop on Algorithms for Modern Massive Data Sets.
I'm presented: "Open Problems in Reproducibility for Massive Datasets"

June 25: International Conference on Machine Learning (top machine learning conference). Giving Workshop Keynote. Title: "Reproducible Research in Machine Learning" (Haifa, Israel – I will also get to hang out with my mom in Haifa, which is cool. I'm also planning a 4 day vacation chilling at a girlfriend's in Jerusalem – I am have never been there so I expect that to be fantastic :))

July 1: Presenting at CERN Colloquium. Title "Reproducible Research, the Scientific Method, and Massive Data" (Geneva)

July 7-11: Kauffman event, Los Angeles. Invited contributor. I wrote a chapter for the book they expect to publish on game-changing rule changes for innovation and growth. I suggested 3. Lots of law bigwigs here..

VISIT JEFFREY

July 29-30: Open Science Summit, Speaker. Title TBA. Berkeley, CA.

July 30-August 1, SciFOO. Invited contributor and speaker.

OR MAYBE VISIT JEFFREY IN HERE

August 12-13 Intellectual Property Scholars, Berkeley CA, Speaker. Title: "Why Copyleft isn't Right for Scientific Code"

Then I am back in NYC and am going to look for a place to live downtown, somewhere in proximity to the 1 train, so probably Chelsea or maybe West Village. I start teaching September 7!

As you can see I am planning to chat with you in person, whenever is convenient :)

* As far as 2) *work* goes:

I am about to take on the machine. Well, here's what I mean. I had an interesting meeting with Eben Moglen, Columbia law prof who wrote the GPL version 3 (near ubiquitously used open license for code), in which he explained that the goal is... Communism: step 1 is using the GPL license – designed to obliterate intellectual property in code. Step 2 is removing intellectual property in computer hardware, and step 3 is removing property rights in bandwidth, ie. the net neutrality movement. If you don't believe me, see http://emoglen.law.columbia.edu/my_pubs/dcm.html Anyway, regardless of the communist stance, this movement's approach is to use a licensing provision called "share alike" which requires any derivative code to come under the exact license of the upstream code. I think this is counter to scientific norms: that we don't dictate terms of use of our scientific outputs, rather knowledge is considered a public good. So I'm writing this up in a paper now, which the pro-communists will despise.

Otherwise I am planning the structure of my role at Columbia. I am thinking of creating a small center, like the Information Society Project I working in last year at Yale, at Columbia. As a faculty member I can apply for grants directly and I think a center is what is needed to bring attention to this issue. There are many issues to research and this would be a nice way to structure them and get students involved.

I'm also slowly shifting my research to include quantitative work on massive data – this is in keeping with my theme of the impact of technology change on science. The statistical methods are rapidly changing to keep pace with changes in the data: think social network data, streaming sensor data, and data at extremely large scale.

Looking forward to chatting more, hope you are well,
Victoria