

## Dan Cohen's Digital Humanities Blog » Blog Archive » When Machines Are The Audience

---

I recently received an email from someone at the Woodrow Wilson Center that began in the following way: “Dear Sir/Madam: I was wondering if you might share the following fellowship opportunity with the members of your list...The Africa Program is pleased to announce that it is now accepting applications...” The email was, of course, tagged as spam by my email software, since it looked suspiciously like what the U.S. Secret Service calls a 419 fraud scheme<sup>[1]</sup>, or a scam where someone (generally from Africa) asks you to send them your bank account information so they can smuggle cash out of their country (the transfer then occurs in the opposite direction, in case you were wondering). Checking the email against a statistical list of high-likelihood spam triggers identified the repeated use of words such as “application,” “generous,” “Africa,” and “award,” as well as the phrases “submitted electronically” and the opening “Dear Sir/Madam.” The email piqued my curiosity because over the past year I’ve started altering some of my email writing to avoid precisely this problem of a “false positive” spam label, e.g., never sending just an attachment with no text (a class spam trigger) and avoiding the use of phrases such as “Hey, you’ve got to look at this.” In other words, I’ve semi-consciously started writing for a new audience: machines. One of the central theories of humanities disciplines such as literature and history is that our subjects write for an audience (or audiences). What happens when machines are part of this audience?

As the Woodrow Wilson Center email shows, the fact that digital text is machine readable suddenly makes the use of specific words problematic, because keyword searches can much more easily uncover these words (and perhaps act on them) than in a world of paper. It would be easy to find, for instance, all of the emails about Monica Lewinsky in the 40 million Clinton White House emails saved by the National Archives because “Lewinsky” is such an unusual word. Flipping that logic around,

if I were currently involved in a White House scandal, I would studiously avoid the use of any identifying keywords (e.g., “Abramoff”) in my email correspondence.

In other cases, this keyword visibility is desirable. For instance, if I were a writer today thinking about my Word files, I would consider including or excluding certain words from each file for future research (either by myself or by others). Indeed, the “smart folder” technology in Apple’s Spotlight search or the upcoming Windows Vista search can automatically group documents based on the presence of a keyword or set of keywords. When people ask me how they can create a virtual network of websites on a historical topic, I often respond by saying that they could include at the bottom of each web page in the network a unique invented string of characters (e.g., “medievalhistorynetwork”). After Google indexes all of the web pages with this string, you could easily create a specialized search engine that scans only these particular sites.

“Machine audience consciousness” has probably already infected many other realms of our writing. Have some other examples? [Let me know<sup>\[2\]</sup>](#) and I’ll post them here.

This entry was posted on Thursday, March 2nd, 2006 at 11:24 am and is filed under [Email<sup>\[3\]</sup>](#), [Search<sup>\[4\]</sup>](#), [Spam<sup>\[5\]</sup>](#), [Text Mining<sup>\[6\]</sup>](#). You can follow any responses to this entry through the [RSS 2.0<sup>\[7\]</sup>](#) feed. You can [leave a response<sup>\[8\]</sup>](#), or [trackback<sup>\[9\]</sup>](#) from your own site.

## References

1. [^ the U.S. Secret Service calls a 419 fraud scheme](#)  
([www.secretservice.gov](http://www.secretservice.gov))
2. [^ Let me know](#) ([www.dancohen.org](http://www.dancohen.org))
3. [^ View all posts in Email](#) ([www.dancohen.org](http://www.dancohen.org))
4. [^ View all posts in Search](#) ([www.dancohen.org](http://www.dancohen.org))
5. [^ View all posts in Spam](#) ([www.dancohen.org](http://www.dancohen.org))

6. [^ View all posts in Text Mining \(www.dancohen.org\)](#)
7. [^ RSS 2.0 \(www.dancohen.org\)](#)
8. [^ leave a response \(www.dancohen.org\)](#)
9. [^ trackback \(www.dancohen.org\)](#)

Excerpted from *Dan Cohen's Digital Humanities Blog » Blog Archive » When Machines Are the Audience*

<http://www.dancohen.org/2006/03/02/when-machines-are-the-audience/>

---

READABILITY — An Arc90 Laboratory Experiment

<http://lab.arc90.com/experiments/readability>